The Application of Protein Mass Spectrometry to the Understanding of Behaviour in *Mus* Species

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Victoria J Lee

October 2015

Acknowledgements

I would like to thank my supervisors Professor Rob Beynon and Professor Jane Hurst for giving me the opportunity to undertake this research and for all their help, guidance and enthusiasm over the past four years.

Many thanks to all the members of CPR past and present for all their help, support and friendship. Special thanks go to – Lynn, for all the chats, laughs and being a smashing friend; and to Lupe, for being a fountain of MUP knowledge, a great mate and for coming to Greggs with me when I needed pie. Also thanks to Stephen for imparting mass spec knowledge, for helping me prepare for conferences and for sharing my sense of humour; you're a top mate, cheers. Thanks to Philip, for all of his technical assistance with instruments. I'd also like to thank the MBE group, who looked after me when I spent a month working with the mice and showed me where to buy cake. Huge thanks goes to Amanda (my maxi - me) for her help, sample collections, laughs and friendship; thanks also goes to Jonathon for allowing me to collaborate with him on the communal nursing study. Thanks to Lee and Hans at Waters for giving me the opportunity to work with them on the multi-omics study.

I would like to thank my wonderful friends – Emily, Jenny, Yvonne and Manda: I couldn't have done this without your constant support and love. Thank you for the fabulous times filled with wine, cocktails and copious amounts of food. Here's to many more.

A special thanks to my best mates, my Mum and Dad. Thank you for your love and everything you have done to help and support me throughout this – from mucking out my horse's stable to delivering chocolate to my house. Much love to Eddie, my equine partner in crime, and Lionel, who has kept me company throughout the writing process, even if it was only so he could get more raisins.

Finally, I would like to thank my partner Joe. Thank you for making me a coffee every morning, for the hugs, for dragging me away from my thesis whenever I needed a break and for making me laugh every single day.

Contents

List of Figures

List of Tables

Appendices

Contributions to publications	An outline of the contributions made to publications.
Supplementary material A	*M. spicilegus de novo* sequencing analysis, additional peptide map and recorded behaviour from behavioural experiments for Chapter 3.
Supplementary material B	Diet preparation and precursor RIA calculation spreadsheets for Chapter 4.
Supplementary material C	Linearity of protein load and ESI-MS response, showing detector saturation at higher protein loads, for Chapter 5.

Supplementary material is located on the disc located in the back of this thesis.

Abbreviations

2D-PAGE	Two dimensional polyacrylamide gel electrophoresis
°C	Degrees centigrade
× *g*	Times gravity
µg	Microgram
µl	Microlitre
A	Alanine
ACN	Acetonitrile
Ala	Alanine
AOB	Accessory olfactory bulb
APEX	Absolute protein expression
AQUA	Absolute quantification
Asp N	Endoprotease Asp N
BLAST	Basic Local Alignment Sequence Tool
BSA	Bovine serum albumin
C	Cysteine
CID	Collision induced dissociation
D	Aspartic acid
Da	Dalton
DTT	Dithiotreitol
E	Glutamic acid
E. coli	Escherichia coli
ESI-MS	Electrospray ionisation - mass spectrometry
ESP	Exocrine gland secreting peptide
ETD	Electron transfer dissociation
F	Phenylalanine
fmol	Femtomole
G	Glycine
GC-MS	Gas chromatography - mass spectrometry
GG	Grueneberg ganglion
Glu C	Endoprotease Glu C
H	Histidine
HCD	Higher-energy collisional dissociation
HCl	Hydrochloric acid
I	Isoleucine
ICAT	Isotope-coded affinity tag
IEF	Isoelectric focusing
Ile	Isoleucine
iTRAQ	Isobaric tag for relative and absolute quantification
K	Lysine
kDa	Kilodalton
kV	Kilovolts
L	Leucine
L	Litre
LC-MS	Liquid chromatography - mass spectrometry
Leu	Leucine
Lys	Lysine
Lys C	Endoprotease Lys C
M	Methionine
M	Molar
MALDI–ToF-MS	Matrix-assisted laser-desorption ionization–time of flight
MaxENT	Maximum entropy modelling
Met	Methionine
mg	Milligram
MHC	Major histocompatibility complex
ml	Millilitre

mm	Millimetre
mM	Millimolar
MOB	Main olfactory bulb
MOE	Main olfactory epithelium
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MUP	Major urinary protein
MWCO	Molecular weight cut-off
My bp	Million years before present
m/z	Mass to charge ratio
N	Asparagine
Native PAGE	Native polyacrylamide gel electrophoresis
nl	Nanolitre
nm	Nanometre
NMR	Nuclear magnetic resonance spectroscopy
N-terminal	Amino-terminal
OR	Olfactory receptor
OSN	Olfactory sensory neuron
P	Proline
PAI	Protein abundance index
Phe	Phenylalanine
pI	Isoelectric point
PMF	Peptide mass fingerprint
pmol	Picomole
ppm	Parts per million
Pr:Cr	Protein:creatinine ratio
PSAQ	Protein standard for absolute quantification
PTM	Post translational modification
Q	Glutamine
QconCAT	Quantification concatamer
Q-peptide	Quantification peptide (from QconCAT)
Q-ToF	Quadrupole/Orthogonal Time of Flight Mass Spectrometer
R	Arginine
rMUP	Recombinant major urinary protein
RO	Reverse osmosis
rpm	Revolutions per minute
S	Serine
SAX	Strong anion exchange chromatography
SBSE	Stir bar sorptive extraction
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC	Stable isotope labelling by/with amino acids in cell culture
SPME	Solid phase microextraction
T	Threonine
TFA	Trifluoroacetic acid
Tris	2-amino-2-hydroxymethyl propane-1, 3-diol
Try	Tryptophan
UPLC	Ultra performance liquid chromatography
V	Valine
V	Volts
V1R	Vomeronasal family 1 receptor
V2R	Vomeronasal family 2 receptor
Val	Valine
VNO	Vomeronasal organ
VSN	Vomeronasal sensory neuron
W	Tryptophan
Y	Tyrosine

Abstract

The Application of Protein Mass Spectrometry to the Understanding of Behaviour in *Mus* Species

The urine of mice contains a large amount of protein, approximately 99% of which are major urinary proteins (MUPs). These 18 – 19 kDa proteins have been shown to have a significant role in chemosignalling. MUPs bind, protect and slowly release the volatile components in scent marks that have been found to elicit various behavioural and physiological responses in mice, including aggression between male mice and the onset of puberty in female mice. MUPs themselves have also been found to be significant in chemical signalling, having roles in modulating identity signalling, attractiveness and aggressive responses.

Whilst the MUPs of the house mouse have been comprehensively characterised, including their roles in social and reproductive behaviour, much less in known regarding the urinary proteins of other closely related *Mus* species and their roles in communication amongst conspecifics. In the first part of this thesis, protein expression in the urine of *Mus spicilegus* was investigated. Whilst *M. spicilegus* are genetically close to the *Mus musculus* subspecies group, their mating patterns and social behaviour are different to those of the house mouse, *Mus musculus domesticus.* Urinary MUPs play a significant part in the social and reproductive behaviours displayed by *M. m. domesticus,* yet little is known regarding the molecular causes of the unique social and reproductive behaviours displayed by *M. spicilegus*. Therefore, this part of the thesis determined whether *M. spicilegus* also invested in MUPs, and by inference, whether these could play a key role in their unusual behaviours. The protein content of the urine of male and female *M. spicilegus* was examined, and mass spectrometry was used to identify four MUPs (three of them male-specific) and characterise their primary structure in the absence of genomic data. Male mice expressed more MUP in their urine than female mice, with MUP expression patterns between different males varying quite significantly. In experiments relating to their sexual and social behaviour, male mice altered their MUP expression upon contact with females, and a link between male MUP output and aggressive behaviour amongst male mice was observed.

As MUPs have roles in modulating a number of behaviours in the house mouse, including kin recognition, the second part of the thesis focuses on communal nursing in the house mouse, where kin recognition appears to play a role in the lactative investment in pups in a communal nest. Previous studies suggest communal nursing

increases reproductive success, but that mutualistic cooperation is higher in nests where females are related. This study involved the metabolic labelling of adult female mice, who were communally nursing, using stable isotope labelled amino acids. The mass spectrometric analysis of MUPs expressed in adult female mice (mothers) urine, proteins in pup stomach contents and proteins in pup tissue samples enabled the confirmation of amino acid label incorporation into nursing mothers, and the determination of label incorporation into pups. Therefore, the investment received by pups from their mothers, in litters where their mothers are related and in litters where their mothers are unrelated was determined, assessing whether any discrimination in investment is evident in relation to relatedness of the female pairs in the nest. Stable isotope labelling strategies and mass spectrometry successfully enabled investment from communally nursing female mice to be tracked in their pups, determining that no females, related or unrelated, discriminated between their own pups and their female partners' when investing. In most cases, however, one female appeared to invest significantly more in the entire communal litter than the other.

Mass spectrometry has been the main tool for the accurate identification and characterisation of MUPs present in scent marks, but MUP quantification has proved more difficult due to the highly homologous nature of these proteins. Previous absolute quantification of MUPs has been based on QconCAT technology, but difficulties arose due to the high sequence similarity between MUP variants. It was therefore considered whether quantification of MUPs could take place at the intact protein level, since intact protein analysis by ESI-MS is already well established for the accurate identification of MUP isoforms. For ESI-MS analysis to be a suitable method for absolute quantification of MUPs, the responses of each individual MUP in ESI-MS analysis were determined using recombinant MUPs, assessing the relationship of MUP concentration and instrument response whilst considering how charge state distribution profiles and sample complexity affected MUP ionisation and instrument response. Despite the homology of MUPs, the differences in their ionisation efficiencies in ESI-MS analysis as part of an equimolar mixture compared to as a single protein means that whilst ESI-MS analysis of MUPs can be useful for the relative quantification of these proteins in urine samples, further preparatory experiments would be required to determine whether ESI-MS analysis of intact proteins could be suitable for absolute quantification.

Chapter 1: Introduction

1.1 Olfactory communication in mice

Communication between animals occurs when one individual's actions provide a signal that changes the behaviour of another (Wiley 1983). Communication in most mammals, including mice, takes place via a mixture of the senses, but is primarily through their sense of smell. Detection of olfactory cues from the environment are essential for avoiding predators, navigating and foraging for food (Doty 1986; Yang and Crawley 2009) but mice also use separate olfactory cues to provide information regarding their sex, social status, health and territories (Beynon *et al*. 2008). These cues, when detected by the recipients' olfactory systems, evoke a number of physiological and psychological responses in the recipient (Bruce 1970).

The term 'pheromones' was first defined by Karlson and Luscher as "substances which are secreted to the outside by an individual and received by a second individual of the same species, in which they release a specific reaction, for example, a definite behaviour or a developmental process" (Karlson and Luscher 1959). Pheromonal communication varies considerably between species, and is key in the regulation of behaviours in mice, such as reproduction, maternal care and aggression (Swnaney and Kaverne 2009). Pheromones can be defined as 'releasers', which trigger immediate, short-lived responses in the receiver; 'primers', which trigger longer-term behavioural or physiological responses; or 'signallers', which convey information about an individual, the responses being dependent on the individuals sending and receiving the signal (Swnaney and Kaverne 2009).

Research into olfactory communication in mammals has primarily focused on the chemical signalling and chemosensory systems of rodents, which has enabled us to gain understanding on the physiological processes and complex psychological behaviours seen in mice, and the different components of the olfactory cues that trigger these various responses.

1.2 Scent marking

Mice introduce chemical signals into their environment mainly in the form of scent marks, enabling an individual to leave information about themselves to conspecifics when they are no longer present (Johnson 1973). Scent marks are usually deposited

via urine, but can also be set down in the form of faeces or specialised scent gland excretions (Brown and MacDonald 1985).

Scent marks deposited by mice contain information regarding the sex, individual identity, social status, health and territories of the individual (Hurst *et al*. 2001; Beynon *et al*. 2008). The location and spatial arrangement of an individual's scent mark define their territory, and help in defending their territory against others in their environment. Dominant male mice will deposit scent marks frequently in order to keep their marks constantly fresh, to defend their territory against intruders and advertise themselves to, and attract, a female mate (Gosling 1982; Hurst 1993; Rich and Hurst 1999; Humphries *et al*. 1999). Dominant male mice will also counter-mark scents deposited by intruders in their territory to re-affirm their status and territorial boundaries, but will not counter-mark their own scent marks or scent marks from another male who is genetically identical to himself (Hurst 1990a; Nevison *et al.* 2003).

1.2.1 Behavioural responses to scent marks

Scent marks, from both male and female mice, can elicit various psychological and physiological responses in conspecifics of the same and opposite sex. The majority of the behavioural responses observed in mice to scent marks relate to their attraction to particular scent marks, in order to find a suitable mate. Generally, post-puberty female mice are more attracted to scent marks from a sexually active male (Hurst 1990(b); Petrulis 2013). In response to male mouse scent marks, reproductively-active adult female mice scent mark in response to a male's scent marking in order to advertise their sexual receptivity (Rich and Hurst 1999). A female prefers scents from male mice they have previously encountered, and appear to develop a preference for the volatiles present in a particular male's scent marks upon contacting that male's urine (Ramm *et al*. 2008). A female's attracted response to male urine is also modulated by the spatial arrangement of scent deposition which indicates that the male is a territory owner (Gosling *et al.* 2001).

In response to female scent marks, male mice produce ultrasonic vibrations (USVs) (Nyby *et al.* 1977). Male urine induces aggressive behaviour in other adult males, but resident males will not attack a castrated male or female intruder (Mugford and Nowell 1971; Mucignat – Caretta *et al*. 2004). A dominant male will also countermark an intruder's scent mark to reaffirm its status and territory.

1.2.2 Physiological responses to scent marks

Several physiological responses to scent marks have been observed in female mice, often to a dominant, reproductively active male's scent. In juvenile females, exposure to these male scents accelerate the onset of puberty, this response being known at the 'Vandenbergh effect' (Vandenbergh 1969). Urine from singly housed females, who are pregnant, lactating or in oestrus, can also cause this response in juvenile females (Drickamer and Hoover 1979). Also in response to a dominant, reproductive male's scent, ovulation in females can be induced; this has been termed the 'Whitten effect' (Whitten *et al.* 1968). Conversely, ovulation can be inhibited in single-sex-group-housed female mice if they are exposed to urine from group-housed females; this is known as the 'Lee-Boot effect' (Van der Lee and Boot 1956). This particular situation can also delay puberty in female mice. (Colby and Vandenbergh 1974). The urine of an unfamiliar male will cause a recently mated female to lose her pregnancy and return to oestrus (known as the 'Bruce effect') (Bruce 1960).

Much less is known about the physiological responses to scent marks by male mice (Koyama 2004), but as with females, it appears that male reproductive physiology is sensitive to chemosignals (Petrulis 2013). For example, puberty in males can also be delayed if exposed to urine of group-housed females, the same as in females (Jemiolo and Novotny 1994). Female scents can also cause dominant (but not subordinate) males to increase sperm production (Petrulis 2013) and to release luteinising hormone (LH) within half an hour of exposure to female urine (Macrides *et al*. 1975; Schulz *et al.* 2009).

The behavioural and physiological responses observed in mice are largely due to the detection of a number of pheromonally active compounds found in scent marks, some of which are volatile, some of which are involatile, and these compounds are explored further in the next section.

1.3 Volatile chemical signals

Research into the chemical components of scent marks has focused on the volatile and non-volatile compounds present, which are associated with signalling the information on species, sex and individual identity of the depositor of the scent mark (Beynon *et al*. 1999). Volatile pheromones are often metabolic by-products, and so the production of these signals requires no extra energy input from the scent mark depositor (Wyatt 2009). Airborne volatile pheromones can draw an animal to investigate a scent mark even when the depositor is no longer in the immediate

vicinity; however, the volatile nature of these compounds means that they are present for a relatively short space of time. A number of volatile pheromones have been identified in mouse urine, the names and structures of which are outlined in Table 1.1.

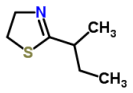
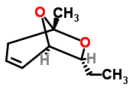
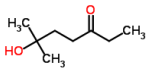
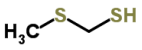
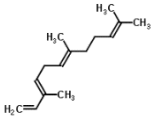
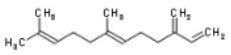
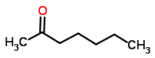
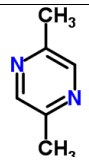
At least four of these volatiles (2-sec-butyl-4,5-dihydrothiazole, 2,3-dehydro-exo-brevicomin, α and β farnesenes) have been associated with social dominance in adult male laboratory mice (Humphries *et al*. 1999). 2-sec-butyl-4,5-dihydrothiazole (thiazole) and 2,3-dehydro-exo-brevicomin (brevicomin) are both volatile but are normally bound to protein in the urine. They are detected by the vomeronasal organ (VNO) of the olfactory system, acting as 'releaser' pheromones on a male recipient and 'primers' on female recipients, and their production is dependent on testosterone (Swaney and Keverne 2009). In male mice, the presence of these compounds in urine causes a persistent aggressive response. If either one of these compounds is mixed into a castrated male's urine and placed on their fur, intact males do not show aggressive behaviour towards the castrated male, but if both compounds were mixed into the castrated urine, this causes males to attack as they would if the urine was from an intact male (Novotny *et al.* 1985; Jemiolo *et al.* 1985). Either one of these volatiles have been found to accelerate the onset of puberty in females, as does 6-hydroxy-6-methyl-3-heptanone (Novotny *et al*. 1999), while a mixture of both thiazole and brevicomin has been found to induce oestrus (Jemiolo *et al*. 1986). Male urine also contains α and β farnesenes, which are produced by the preputial gland and released into the deposited urine. These volatiles are detected by the VNO and are also known to attract females, as well as signalling dominance to other males, inhibiting investigation and countermarking in subordinate male mice (Novotny *et al.* 1990; Jemiolo *et al*. 1991, Jemiolo *et al*. 1992). Also present in male urine, (methylthio)-methanethiol, detected by the main olfactory epithelium (MOE), is an attractant to female mice (Lin *et al.* 2005).

Present in female urine, 2.5-dimethylpyrazine delays the onset of puberty in juvenile female mice (Ma *et al*. 1998), and 2-heptanone has been seen to extend the length of oestrus in females (Jemiolo *et al*. 1989). Both of these volatiles are detected in the VNO.

1.4 Non-volatile chemical signals

While volatile signals draw an animal to investigate a scent mark, an animal must physically contact the scent mark in order to detect the non-volatile components. Non-volatile chemical signals offer information about the individuality of the signaller

Table 1.1 Volatile compounds found in mouse urine and their structures (obtained from www.chemspider.com)

Volatile	Structure	Origin
2-sec-butyl-4,5-dihydrothiazole		Male urine
2,3-dehydro-exo-brevicomin		Male urine
6-hydroxy-6-methyl-3-heptanone		Male urine
(methylthio)-methanethiol		Male urine
α farnesene		Male preputial gland
β farnesene		Male preputial gland
2-heptanone		Female urine
2.5-dimethylpyrazine		Female urine

and the genetic relationship between the scent mark depositor and the detecting animal, information which is important for mate choice decisions and post-mating behaviours such as kin recognition (Hurst *et al*. 2001; Boehm and Zufall 2006; Cheetham *et al*. 2007; Sherborne *et al*. 2007; Thom *et al*. 2008). Unlike volatiles, which are only present for a relatively short space of time, non-volatiles are much more stable and remain in the scent mark for a longer period of time, allowing the individual identity of the depositor to be detected by another animal as the scent mark ages. The non-volatiles identified as significant in mouse social behaviour and interactions are exocrine gland-secreting peptides (ESPs), major histocompatibility complex (MHC) peptides and major urinary proteins (MUPs).

ESPs

Non-volatile chemical signals have been identified in the tear fluid of mice; a total of 38 ESP family genes have been identified in mice, clustered on chromosome 17, 14 of which are pseudogenes (Kimoto *et al.* 2005, 2007). ESP genes encode peptides that are secreted into tear fluid via the extraorbital lachrymal gland, or from the harderian gland or submaxilliary gland, and appear to elicit a response in the vomeronasal epithelium via direct contact (Kimoto *et al*. 2007). To date, one peptide has been found to be male-specific, another has been found to be female-specific, and great variation in ESP expression has been observed between different mouse strains (Kimoto *et al.* 2007). This variation between sexes and strains, plus the responses to the peptides in the VNO, suggests that ESPs may convey information regarding sex and individuality amongst conspecifics (Kimoto *et al.* 2007).

MHC peptides

The genes of MHC are highly polymorphic loci that encode cell surface glycoproteins, that are membrane-bound and anchored in the lipid bilayer (class I and class II molecules), that bind peptides for T lymphocyte–mediated immune recognition of pathogens (Klein 1986; Novotny *et al*. 2007; Tirindelli *et al*. 2009). The genetic diversity of MHC genes is reflected into structurally diverse peptide binding regions of MHC molecules (Apanius *et al*. 1997; Edwards and Hedrick 1998), meaning that different MHC molecules bind to different peptides (Tirindelli *et al.* 2009). These peptides, typically nine amino acid residues in length, are presented by the MHC molecules at the cell surface (Thompson *et al*. 2007). As the MHC is cleared from the cell, the peptide binding region of the MHC molecule releases the peptide into extracellular fluid, and the peptides or their fragments are then excreted in urine and other secretions (Singh *et al*. 1987). The structures of these peptide ligands, and

therefore the structures of the MHC molecules, provide a unique molecular identity signature for each individual (Leinders-Zufall *et al*. 2004).

MHC genes are crucial in influencing resistance and susceptibility to infectious and autoimmune diseases, with the class I and class II molecules initiating specific immune responses against pathogens and parasites (Klein 1986; Penn 2002). Laboratory mice prefer to mate with mice that possess a dissimilar MHC type to its own (Yamazaki *et al*. 1976; Penn and Potts 1998a), and it is thought that the aim may be to produce offspring with increased MHC heterozygosity, resulting in increased resistance to disease (Potts and Wakeland 1993; Kurtz *et al* 2006). Further studies have presented evidence for the effects of MHC genes in mating preference (Penn and Potts 1998a; 1999; Beauchamp and Yamazaki 2003). Whilst it has been suggested that these MHC effects serve to increase MHC heterozygosity and increased resistance to infectious diseases in offspring (Penn and Potts 1999), evidence for this immunological advantage to mate selection is extremely limited, with one study suggesting that MHC heterozygosity appears to provide no immunological advantage (Ilmonen *et al*. 2007). Another hypothesis was that MHC similarity, used to recognise kin, served to avoid inbreeding (Brown and Eklund 1994). The MHC genotype has again been linked to recognition when considering the 'Bruce effect' seen in female mice (Bruce 1959). Leinders-Zufall *et al.* demonstrated that pregnancy block was brought about in female mice by exposing them to MHC class I peptides of a different mouse strain (Leinders-Zufall *et al*. 2004).

MHC diversity between individual mice has also been proposed as a primary source of scents used in individual identity signalling, however, since native MHC peptides have not been identified in mouse urine, it remains uncertain whether they have role in individual recognition. Hurst *et al*. discovered that MHC-associated odours were not necessary nor sufficient for scent owner recognition amongst male mice (Hurst *et al.* 2005), and further studies confirmed that female recognition of scents from male mice depends on a difference in MUP type, and not MHC type (Cheetham *et al.* 2007). It has been demonstrated that wild mice breeding in a semi-natural environment showed no avoidance of mates with the same MHC genotype when genome-wide similarity was controlled, but by contrast, sharing of both MUP haplotypes had a highly significant effect on the likelihood of successful mating, with the strong deficit explaining inbreeding avoidance and kin recognition (Sherborne *et al*. 2007). Whilst it has been demonstrated that mice are able to discriminate between urinary MHC odours, suggesting that MHC type can influence the volatile profile of mouse urine

(Yamaguchi *et al.* 1981; Yamazaki *et al.* 1983; Penn and Potts 1998b; Carroll *et al.* 2002), individual recognition and assessment of genetic heterozygosity in house mice is mediated by major urinary proteins (Hurst *et al*. 2001; Cheetham *et al*. 2007; Thom *et al*. 2008).

1.4.1 Major Urinary Proteins (MUPs)

Mouse urine contains a large amount of protein – in humans, an excess of protein in urine would be of medical concern, as it is usually indicative of impaired kidney function or diabetes, but rodents excrete a significant amount of protein (up to 20 mg/ml per day) whilst maintaining renal function (Gosling *et al*. 2000). In mice, 99% of the proteins excreted in urine are major urinary proteins (MUPs). Mouse MUPs are encoded by a multigene cluster located on chromosome 4, and are mainly synthesised in the liver before being directly excreted in urine (Krauter *et al.* 1982). These proteins are 18 – 20 kDa monomeric proteins that have been shown to have a significant role in chemosignalling (Beynon and Hurst 2003), and are part of the lipocalin superfamily.

Lipocalins are a large family of small, secreted proteins that exhibit great sequence diversity, but share short, conserved motifs that indicate family membership (Flower *et al.* 1993). Despite their sequence dissimilarity, their crystal structures are highly conserved, comprising of a single eight-stranded antiparallel β-barrel which encloses an internal ligand binding site (Flower 1996). The amino acid sequence diversity in this site gives this family of proteins the ability to bind to a wide range of small hydrophobic molecules, giving rise to various biological functions such as prostaglandin synthesis, retinol transport and, importantly in mice, olfaction and pheromone transport (Flower 1996). X-ray crystallography has allowed the structures of mouse MUPs to be determined (Bocskei *et al.* 1991; Bocskei *et al.* 1992; Lucke *et al.* 1999; Timm *et al.* 2001), showing they have the characteristic lipocalin structure, with the eight-stranded β-sheet barrel enclosing the ligand binding site which contains numerous hydrophobic amino acid residues (Figure 1.1). The binding site contains the characteristic tryptophan residue (Try 19) in the centre, which is conserved amongst all lipocalins (Flower *et al*. 1993).

MUP genetics

Sequencing of the C57BL/6 laboratory mouse strain genome allowed the targeted sequencing, annotation, phylogenetic and genomic analysis of the MUP genes in this particular strain by Mudge *et al*. in 2008, work which was built upon by Logan *et al*.

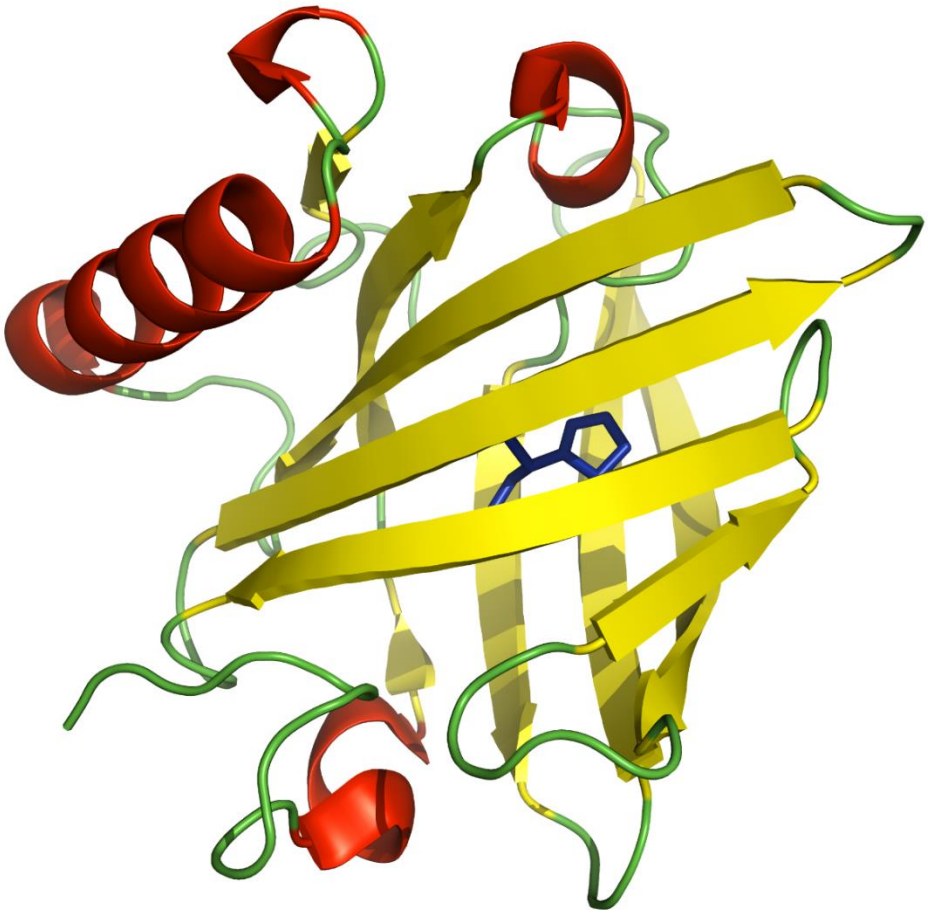


Figure 1.1 The tertiary structure of mouse MUP 1 with ligand.
A ribbon diagram of the 3D structure of mouse MUP 1 along with secondary domains. In yellow is the β sheet forming the β barrel, in red is the the α helix. At the centre of the cavity is the male specific ligand 2-sec-butyl -4, 5 dihydrothiazole, which is coloured blue. The tertiary structure and ligand was solved by Timm *et al*. 2001, and this figure was generated using PyMOL molecular visualisation software (Schrodinger, Inc).

2008. Manual annotation of the MUP cluster allowed Mudge *et al.* to identify 19 functional MUP genes and 19 pseudogenes, with Logan *et al*. later finding evidence for 21 functional and 21 pseudogenes. The MUP cluster was then separated into three groups based on phylogenetic analysis – one group contained highly homologous functional MUP genes, another contained only pseudogenes, and the final group contained functional and pseudogenes that were more divergent and shared lower homology with the other MUP genes. Mudge *et al*. 2008 localised these groups within the MUP locus to two areas, named as the 'central' and 'peripheral' regions. The central region, containing 15 functional MUP genes and 16 pseudogenes, is flanked at either end by the peripheral region, which contains 6 functional MUP genes and 5 pseudogenes (Figure 1.2) (Mudge *et al*. 2008; Logan *et al*. 2008). The peripheral genes are those that share less sequence homology (Figure 1.3), and the oldest divergence for the functional genes in this region is thought to be around 11.2 – 22.4 million years ago, whilst the functional genes in the central region (which share high sequence homology (Figure 1.3)) are thought to be as a result of divergence from a peripheral MUP gene, with the oldest divergence event for these genes estimated at a much more recent 1.2 – 2.4 million years ago (Mudge *et al.* 2008).

MUP ligand binding

Mouse MUPs have been found to bind a number of volatile components identified as having pheromonal activity in their hydrophobic internal ligand binding site with some specificity (Robertson *et al*. 1993; Armstrong *et al*. 2005). Separation of mouse urine into high and low molecular weight fractions using dialysis and/or chromatography allowed the association between MUPs and the following pheromonal compounds to be identified - 2-sec-butyl-4,5-dihydrothiazole, 2,3-dehydro-exobrevicomin and 6-hydroxy-6-methyl-3-heptanone (Bacchini *et al*. 1992; Robertson *et al*. 1993; Novotny *et al.* 1999). As discussed previously, the volatile nature of these pheromones means that they only remain in a deposited scent mark for a very short period of time. However, when these pheromonal ligands are bound to MUPs, their release into the environment is delayed and their lifetimes are extended, remaining detectable by conspecifics up to 24 hours after being deposited (Hurst *et al*. 1998; Humphries *et al*. 1999). If these ligands are not bound to MUPs, it takes a matter of minutes for them to be lost to the surrounding environment (Robertson *et al*. 2001).

The interactions of MUPs and the pheromones 2-sec-butyl-4,5-dihydrothiazole and 6-hydroxy-6-methyl-3-heptanone have been mapped using NMR analysis,

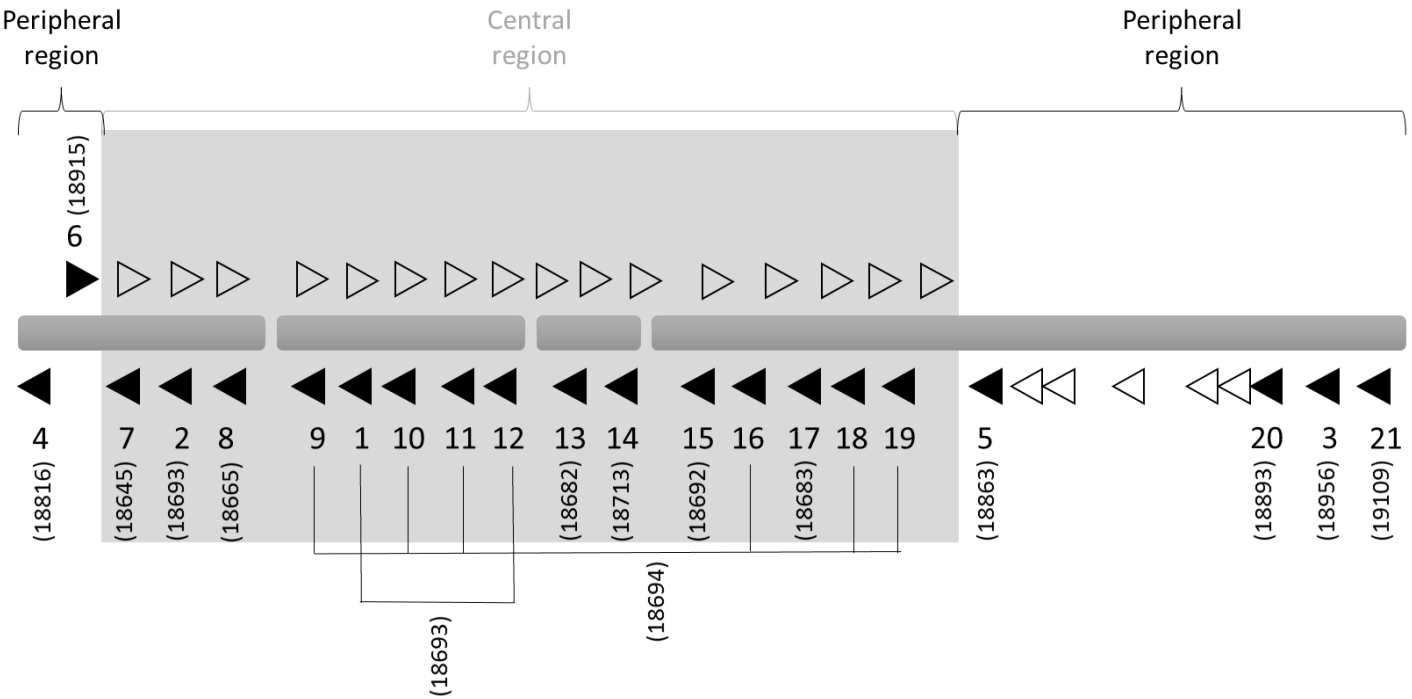


Figure 1.2 C57BL/6 mouse MUP gene cluster.
The central region of the MUP gene cluster is shaded in grey, and the peripheral regions either side in white. All predicted central MUP genes (black triangles) are positioned on the reverse strand, and the central pseudogenes (grey triangles) are positioned on the forward strand. All predicted peripheral MUP genes except MUP 6 (black triangles) are positioned on the reverse strand, along with the peripheral pseudogenes (white triangles). The mature MUP masses are written alongside the black triangles (adjusted for the loss of the signal peptide and the presence of a disulphide bond). Adapted from Mudge *et al.* 2008.

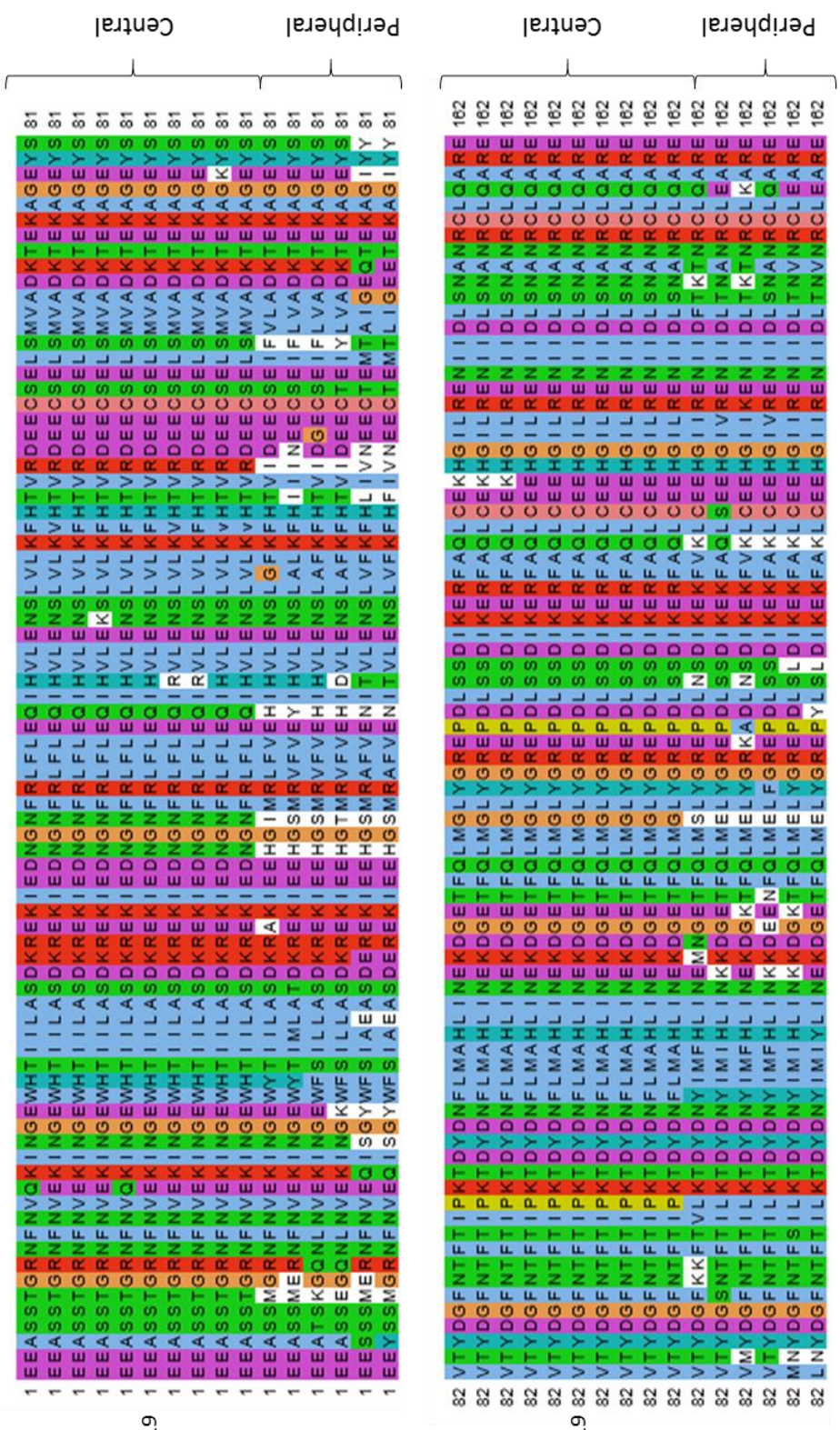


Figure 1.3 Sequence homology of MUPs.
All mature MUP sequences were aligned using Clustal Omega (www.ebi.ac.uk) and the multiple alignment was viewed using Jalview. Residues were coloured using the Clustal option to highlight sequence differences. The aligned sequences are grouped into central and peripheral MUPs.

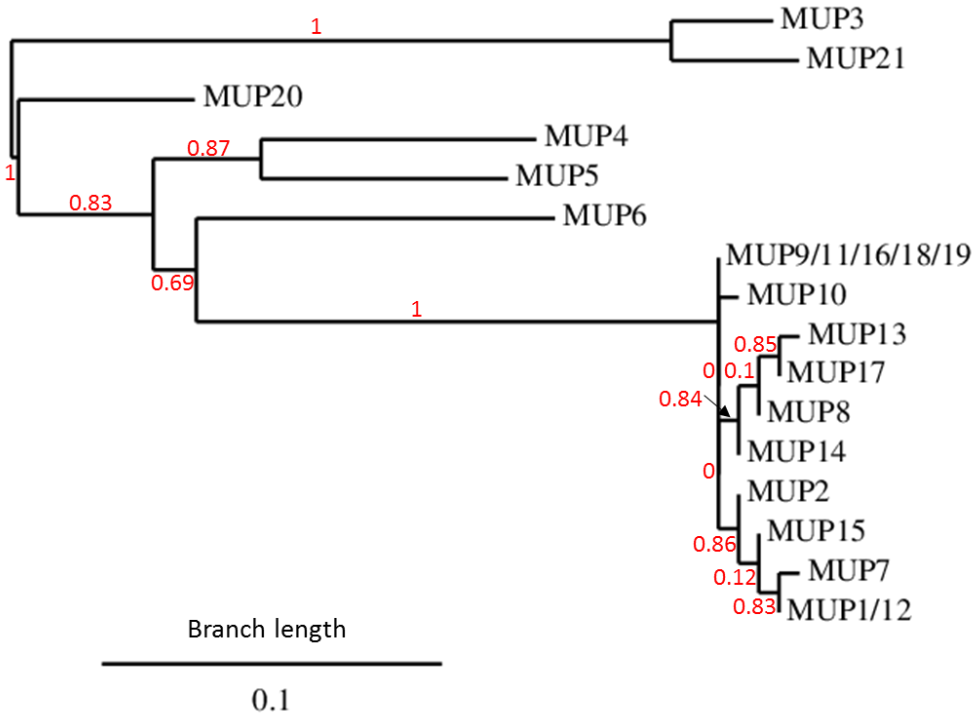


Figure 1.4 Phylogentic tree of mouse MUPs.
Phylogenetic analysis was performed using the 'One Click' mode at www.phylogeny.fr (Dereeper *et al.* 2008). Sequences were aligned using MUSCLE (v3.8.31), configured for highest accuracy. Gblocks (v0.91b) was used to remove any ambiguous regions. The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT). Graphical representation of the tree was performed with TreeDyn (v198.3). Bootstrap values are labelled in red.

characterised using thermodynamic analysis and observed using x-ray crystallography (Bockskei *et al*. 1992; Zidek *et al*. 1999; Timm *et al*. 2001; Sharrow *et al*. 2003). Both pheromonal ligands have been seen to bind to within the hydrophobic MUP ligand binding site at one end of the β-barrel, formed by the side chains of Phe56, Leu58, Leu60,Ile63, Leu72, Phe 74, Met87, Val100, Tyr102, Phe108, Ala121, Leu123, Leu134, and Tyr138 (Timm *et al*. 2001). The precise orientation of the ligand binding has also been established, with a water-mediated hydrogen bond to the 2-sec-butyl-4,5-dihydrothiazole nitrogen, and the ketone oxygen group in 6-hydroxy-6-methyl-3-heptanone (Timm *et al*. 2001). The fact that the MUP hydrophobic binding site is completely enclosed by side chains means that the way in which ligands reach the binding site remains undetermined, but could be through the MUP undertaking large conformational changes (Lucke *et al*. 1999). Through NMR relaxation experiments, it was discovered that the backbone flexibility of the MUP increases as it binds to 2-sec-butyl-4,5-dihydrothiazole, enabling access of the ligand to the MUP's binding site whilst stabilising the MUP-ligand complex (Zidek *et al*. 1999).

The specificity of ligand binding by MUP isoforms was identified by analysis of MUP containing fractions from anion-exchange chromatography (Robertson *et al*. 1993; Armstrong *et al*. 2005), with a study using fluorescent probes supporting this. Darwish Marie *et al*. (2001) proved that the differences in amino acid composition in the ligand binding sites of different MUP isoforms resulted in a decreased ligand binding affinity, causing a decreased fluorescence yield for the probe.

More recently, a male-specific MUP with the molecular weight of 18,893 Da (known as 'darcin') has been found to bind to more 2-sec-butyl-4,5-dihydrothiazole than other MUP isoforms – Armstrong *et al.* (2005) demonstrated that in a C57BL/6 male mouse urine sample, darcin (making up about 13% of the urinary protein concentration) bound to over 40% of the 2-sec-butyl-4,5-dihydrothiazole present in the urine, whilst the other 87% of urinary protein bound to less than 40% of the 2-sec-butyl-4,5-dihydrothiazole. As well as increased ligand binding affinity, darcin was found to bind to 2-sec-butyl-4,5-dihydrothiazole more tightly and release the volatile component more slowly (Armstrong *et al*. 2005; Roberts *et al*. 2010). Behavioural studies have shown that female mice are more attracted to darcin than other MUPs in male urine, and showed no preference between the native (ligand-bound) darcin and its recombinant form, suggesting that darcin itself may also have a role in pheromonal communication (Roberts *et al*. 2010), as well as, like other MUP isoforms, the role in

transporting, protecting and extending the lifetime of the volatile pheromones present in urinary scent marks.

MUP expression

Urinary MUPs are primarily synthesised in the liver and are excreted into urine, avoiding glomerular filtration due to their small size and monomeric nature. The synthesis of MUPs is hormone-controlled; growth hormone, thyroxine, glucocorticoids and insulin all play a role in normal MUP synthesis (Knopf *et al*. 1983; Spiegelberg *et al*. 1988; Johnson *et al*. 1995), whilst androgens have been found to actively induce MUP synthesis – the administration of testosterone to female or castrated male mice increases MUP mRNA levels to those seen in intact male mice (Ruemke and Thung 1964). MUPs are synthesised with a signal peptide, 19 amino acids in length, which is cleaved from the rest of the protein prior to excretion into the bloodstream (Finlayson *et al.* 1965). MUPs have also been detected in the salivary, lachrymal and mammary glands, as well as in nasal tissues (Shaw *et al*. 1983; Shahan *et al.* 1987; Shi *et al*. 1989; Logan *et al*. 2008). Five of the 21 identified functional MUPs (4 central, 1 peripheral) are known to be expressed in the submaxillary glands of the laboratory strain C57BL/6 (Logan *et al*. 2008).

Several different MUP isoforms have been identified in urine, and although many are highly homologous, their differing masses have allowed the majority of urinary MUPs to be separated and identified using isoelectric focusing (IEF) and mass spectrometry (Robertson *et al.* 1996, 1997; Beynon *et al.* 2002; Cheetham *et al.* 2009; Mudge *et al.* 2008). In laboratory mouse strains, adult males excrete around 10 – 20 mg/ml of protein in their urine, with adult females generally excreting significantly less (Cheetham *et al.* 2009). Laboratory mice are inbred and so all mice of the same sex are genetically identical and are homozygous for their MUP genes. Therefore, mice of the same sex and strain have practically identical MUP expression profiles. In wild mice, both adult male and females excrete around three times as much protein as their laboratory counterparts (Beynon and Hurst 2004), and their MUP expression profiles are far more complex, with significant variation in the profiles of unrelated individuals (Robertson *et al*. 1997; Beynon *et al*. 2002). This variation is due to wild mice inheriting different MUP haplotypes from their parents, and maintained by the avoidance of inbreeding (Sherborne *et al*. 2007). It is this variation in MUP profiles which provides the genetic identity of an individual, as the expressed MUP profile provides the MUP genotype information to a conspecific. Various behavioural experiments have confirmed this role of MUPs, with males countermarking a scent

containing a different MUP profile to its own, but not one with the same MUP profile (Hurst *et al*. 2001). Also, the addition of a recombinant MUP to a male's scent mark, thus changing the MUP expression profile, causes that animal to countermark his own scent mark (Hurst *et al*. 2001).

Roles of MUPs in behaviour

As well as their roles in the binding, protection and release of the volatile components found to elicit behavioural responses in mice, MUPs themselves have roles in modulating identity signalling, attraction and aggressive responses in mice. The countermarking study by Hurst *et al*. 2001 suggests the actual components of urine, rather than just scent, are used by mice to evaluate and respond to the scent mark, with recognition depending on the differences of the male's MUP type. The male-specific MUP darcin, when bound to no pheromonal ligands, attracts female mice, indicating its role in pheromonal communication (Roberts *et al*. 2010). MUPs also appear to cause aggression between male mice, with ligand-free MUPs inducing male aggression and partially activating vomeronasal neurons (Chamero *et al*. 2007).

1.5 Pheromone detection

For chemical signals to elicit behavioural and physiological responses in mice, they must be received and processed by their olfactory systems. The olfactory systems that mice chiefly use for the detection and processing of airborne, volatile odours and non-volatile pheromones are the main olfactory system (MOS) and the accessory olfactory system (AOS). The MOE (part of the MOS) is located at the end of the nasal cavity, and the VNO (part of the AOS) is based in the vomer (Figure 1.5). The MOE and VNO detect chemical signals and process the information in separate neural pathways (Swaney and Kaverne 2009). The two systems converge at the level of the amygdala (Swanson and Petrovich 1998), but the segregation of the MOS and AOS and their functional differences has resulted in the basic view that the MOE is responsible for detecting volatile signals and the VNO detects non-volatile signals (Buck 2000). However, there is in fact overlap in their roles of chemosignal detection – the MOE has been shown to be activated by non-volatile peptides (Spehr *et al*. 2006), and the VNO has been seen to respond to volatile odours (Xu *et al.* 2005). Despite this overlap, each system appears to mediate different behavioural responses (Restrepo *et al*. 2004; Spehr *et al*. 2006).

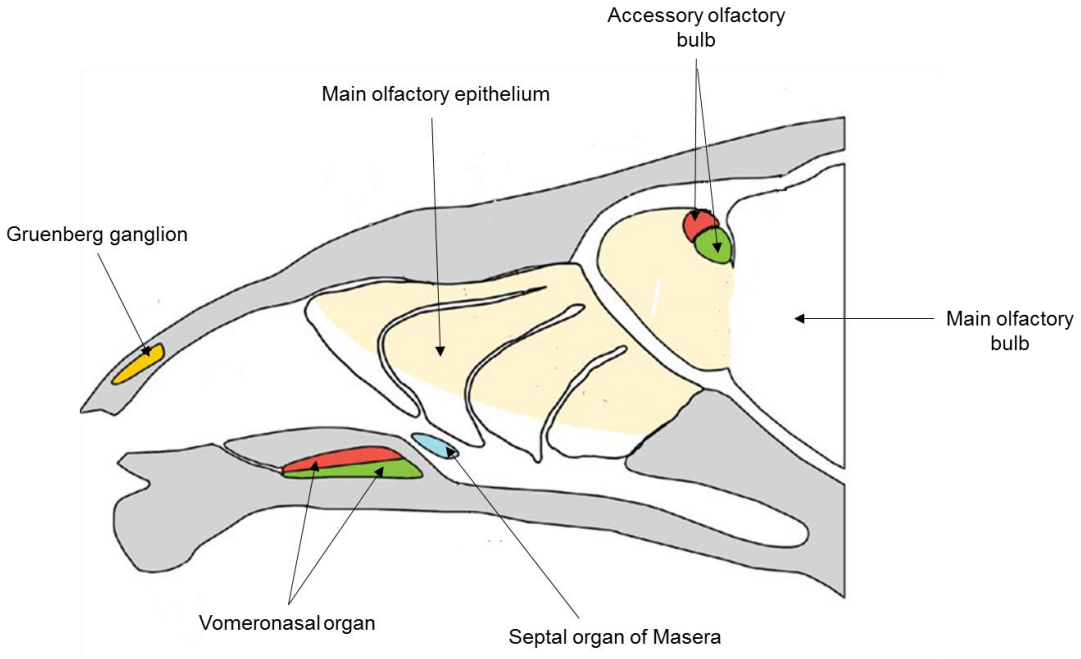


Figure 1.5 Anatomical organisation of the mouse olfactory system.
The location of the chemosensory subsystems found in the nose of a mouse, including the Gruenberg ganglion (GG), the vomeronasal organ (VNO), the main olfactory epithelium (MOE), the accessory olfactory bulb (AOB) and main olfactory bulb (MOB). Adapted from Brennan and Zufall 2006.

1.5.1 Main olfactory system

The MOE of the main olfactory system generally detects volatile, airborne odours, since it is the only one of the two detection systems that has the appropriate morphology and complexity of sensory receptors, plus, sequencing analysis shows that the MOE receptor gene sequences are well conserved across vertebrates (Swaney and Keverne 2009). This suggests that MOE receptors broadly function to detect and process environmental cues rather than species–specific pheromonal cues (Swaney and Keverne 2009).

In the MOS, the MOE is located at the posterior end of the nasal cavity and consists mainly of olfactory sensory neurons (OSNs). The bipolar OSNs project their axons into the main olfactory bulb (MOB), which then projects nerve fibres to the olfactory cortex (Ma 2010). These nerve fibres are then projected from the MOB to the higher sensory centres. Each OSN represents a different olfactory receptor type, of which mice have around 1300 (Zhang and Firestein 2002). These olfactory receptors are from the G-protein-coupled receptor (GPCR) superfamily, and each has a different amino acid sequence, enabling them to bind to a wide range of odourants with different affinities (Zhang and Firestein 2002). Whilst OSNs appear uniform in their morphology, they can be divided into subpopulations based on their chemoreceptors or signal transduction machineries (Ma 2010). These include guanylyl cyclase type–D sensory neurons, which may serve as carbon dioxide sensors and detectors for natriuretic peptides and components of urine; trace amine–associated receptor expressing neurons, which are involved in the recognition of volatile amines found in urine and so might be involved in receiving cues related to sexuality and fear (Liberles and Buck 2006); and transient receptor potential channel expressing cells, which in mammals are grouped into six different subfamilies, some of which have roles in the detection of semiochemicals (Lin *et al.* 2007; Ma 2010).

Sexual behaviour in male mice has been found to be dependent on a functioning MOE, with ablation of the MOE resulting in disrupted investigative and copulatory behaviour (Keller *et al*. 2006). The attraction of female mice to the male pheromone (methylthio)methanethiol has been found to be as a result of mediation by the MOS (Lin *et al*. 2005). As well as the detection and processing of volatile chemical signals, calcium imaging studies have shown that sensory neurons in the MOE respond to MHC peptides (Spehr *et al*. 2006), which supports the more recent view that the MOB and AOB do not operate in mutually exclusive sensory domians (Brennan and Kendrick. 2006).

1.5.2 Accessory olfactory system

The VNO of the accessory olfactory system mostly detects pheromones and so plays a more important role in species-specific communication, a theory supported by the fact that species-specific VNO receptor genes exist, indicating more rapid evolution of the VNO receptor repertoire (Swaney and Keverne, 2009). Alongside the detection of non-volatile components, the AOS has also been seen to respond to certain volatile odours – the mouse pheromone 2–heptanone (found in female urine) and general urine odour were found to elicit significant responses in the accessory olfactory bulb (AOB) (Xu *et al*. 2005).

In the AOS, the VNO is located at the base of the septum which is tubular in shape, blind-ended and filled with fluid, opening into the basal part of the nasal cavity (Droving and Trotier 1998). Similar to the MOE, the VNO contains sensory neurons which project axons, but into the accessory olfactory bulb (AOB) (Keverne 1999). Generally, mammals are seen to curl their top lips back in order to maximise vomeronasal exposure when physically contacting sent marks (known as the Flehman response), causing semiochemicals to be pumped into the lumen of the VNO (Meredith 1994; Swaney and Keverne 2009). These semiochemicals bind to receptors based on the surface of the sensory neurons, triggering them to project into the AOB.

In rodents, the VNO sensory neurons possess two different types of receptors that belong to the G-protein-coupled receptor seven-transmembrane protein superfamily – vomeronasal type 1 and type 2 receptors (V1Rs and V2Rs, respectively) (Dulac and Axel 1995; Herrada and Dulac 1997; Matsunami and Buck 1997; Ryba and Tirindelli 1997). V1Rs and V2Rs are different to one another in terms of sequence, location in the VNO, signalling proteins, and stimuli response. V1Rs are expressed in the apical part of the VNO, binding to the $G\alpha_{i2}$-protein, and the sensory neurons which express V1Rs project only to the anterior AOB (Keverne 1999; Halpern and Martinez-Marcos 2003). V1R-expressing sensory neurons respond to the male specific mouse pheromones 2-sec-butyl-4,5-dihydrothiazole and 2,3-dehydro-exobrevicomin (Leinders-Zufall *et al*. 2000). V2Rs are expressed in the basal part of the VNO, binding to the $G\alpha_{0}$-protein and the sensory neurons expressing V2Rs have been found to bind to MHC class 1 proteins, projecting only to the posterior AOB (Keverne 1999; Halpern and Martinez-Marcos 2003; Leinders-Zufall *et al*. 2004).

Activation of the mouse VNO in terms of behavioural response to pheromones has been studied in some detail, indicating that the VNO plays a significant role in sexual

behaviour in male mice and maternal aggression in female mice (Del Punta *et al*. 2002), in aggressive behaviour between male mice (Stowers *et al*. 2002), the pregnancy block effect (Kelliher *et al*. 2006), copulatory behaviour (Keller *et al*. 2006), the suppression of oestrus in group housed female mice and the induction of puberty in females exposed to male odours (Keverne 1983).

1.6 Discovery, identification and quantification of scent mark components

The approaches for the discovery, identification and quantification of the volatile and non-volatile components of scent marks are multidisciplinary, but primarily involve mass spectrometry. Once a pheromonally active compound has been identified and characterised, synthetic versions can be produced and used in behavioural experiments to confirm their roles in the physiological and psychological responses seen in, for example, mice (Novotny 2003).

In mammals, pheromones are typically embedded in a complex biological matrix, such as urine, requiring the first step in the discovery of volatiles to involve separation of the volatiles of interest from the rest of the sample. Volatiles are extracted from a sample or the headspace above the sample in a sealed container using a solvent, or utilising solid phase microextraction (SPME)/stir bar sorptive extraction (SBSE) techniques before being analysed using capillary gas chromatography coupled with mass spectrometry (GC-MS) (Robertson *et al*. 1993; Soini *et al.* 2005). Since pheromone production is known to be affected by hormones, the metabolic profiles of volatiles of animals in different behavioural or endocrine situations can be assessed by GC-MS, giving an insight into the compounds of interest in these situations (Novotny 2003). This information can then allow more detailed analysis of the pheromones of interest and the conduction of the relevant behavioural experiments. This particular approach has allowed the identification and characterisation of the volatile signals confirmed to be responsible for puberty induction and delay, oestrus induction, attraction, aggression and dominance in mice (Jemiolo *et al*. 1985, 1986, 1991; Ma *et al*. 1999; Novotny *et al*. 1985, 1986, 1990,1999).

The term ‘proteomics’ was coined by Wilkins *et al.* 1996, and advanced proteomic analysis has been instrumental in the identification, characterisation and quantification of the non-volatile protein components of scent marks, even in the absence of genomic data. The use of a wide range of mass spectrometric techniques, including matrix assisted laser desorption ionisation–time of flight mass spectrometry (MALDI-ToF-MS), electrospray ionisation–mass spectrometry (ESI-MS), liquid chromatography coupled with mass spectrometry (LC-MS) and liquid

chromatography with tandem mass spectrometry (LC-MS/MS) has allowed proteins in scent marks to be identified and characterised in terms of their molecular weight, peptide mass fingerprints and amino acid sequences. Chromatographic techniques have enabled the isolation of proteins of interest from complex matrices, and the use of recombinant proteins of interest can be used in behavioural experiments to provide a further insight into their roles in communication. The remainder of this section shall focus on the biochemical methods commonly utilised for the identification and characterisation of major urinary proteins, which have proven to be of great interest in mouse olfactory communication.

1.6.1 Discovery of MUPs in a scent mark

A commonly used method for the discovery of MUPs in a scent mark is by analysing the sample using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). SDS, an anionic detergent, linearises and negatively charges proteins in the sample, allowing them to be separated on the polyacrylamide gel according to their molecular weight. Staining of the gel post-electrophoresis allows the protein content of the sample to be visualised by eye, and the presence of MUPs in the sample would usually result in a band(s) at approximately 20 kDa on the gel (the use of a molecular weight marker on the gel allows the location of ~ 20 kDa proteins to be determined). The MUP darcin, with a mass of 18,893 Da, however, is seen at around 16 kDa on a gel – even after treatment with SDS, darcin does not completely linearise, indicating that protein folding and shape can also affect the migration of proteins during SDS-PAGE analysis (Armstrong *et al*. 2005; Phelan *et al*. 2014). The number of protein bands seen on a gel gives an indication of the protein complexity of a scent mark, whilst the density of the bands gives an approximate indication of the abundance of that protein. In mouse urine (particularly male), a high concentration of a number of different MUPs gives a large, dense band at around 20 kDa, meaning initial identification of the number of MUPs and their relative abundances cannot be determined. Methods such as native PAGE (non-denaturing, meaning proteins can be separated according to folding and shape as well as molecular weight) and 2D-PAGE (where proteins are first separated by charge before being separated again by their molecular weight) allow proteins such as MUPs to be visualised more effectively prior to further biochemical analysis.

1.6.2 Identification of MUPs in a scent mark

Whilst SDS-PAGE analysis can give an approximation of the molecular weights of proteins in a scent mark, mass spectrometry is the primary tool for accurate identification of proteins. An accurate molecular weight for proteins in a sample can be determined via electrospray ionisation mass spectrometry (ESI-MS). Electrospray ionisation is a 'soft' ionisation technique, meaning very little fragmentation occurs during ionisation and so allows the formation of gas phase molecular ions. ESI can also produce multiply charged ions, which extends the mass range of the analyser to accommodate the higher orders of magnitude observed in proteins (Ho *et al*. 2003; Pitt 2009). Emitted ions are accelerated into the mass analyser and are subjected to mass spectrometric analysis, where proteins will exhibit their multiply-charged ions in a cluster (Ho *et al.* 2003). The number of charges on a protein molecule will depend on its molecular weight and the number of accessible basic sites (Ho *et al*. 2003). The molecular weight of the protein(s) can be calculated from the observed protein envelope using software provided by the mass spectrometer manufacturer, which can process and transform the raw mass spectrum to a true mass scale and determine the molecular weight (Armstrong *et al.* 2005). This is an important step in identifying MUPs in a scent mark, as calculated molecular weights can be compared and matched to the known MUP molecular weights to confirm their presence.

To identify the protein bands observed in SDS-PAGE analysis, parts of the bands can be excised before being digested with a protease to create peptide fragments suitable for mass spectrometric analysis. A protease hydrolyses the bonds between amino acids, and the use of different specific proteases (such as Lys-C, which only hydrolyses at the carboxyl side of lysine residues, and Glu-C which specifically hydrolyses at the carboxyl side of aspartic or glutamic acid residues) for proteolysis creates different peptide mass patterns. The peptides resulting from protein digestion can be subjected to mass spectrometric analysis to create what is known as a 'peptide mass fingerprint' (PMF), which can then be searched against a database containing known PMFs to help determine the identity of the protein (Perkins *et al*. 1999). MALDI-ToF-MS is a popular method for accurately measuring the m/z of each peptide and generating peptide mass fingerprints. Again, MALDI is a soft ionisation technique, resulting in no undesired fragmentation of peptides and it usually produces singly-charged ions, generating relatively simple mass spectra suitable for searching against a peptide mass fingerprint database. Searching against a relevant PMF database involves matching the experimental peptide masses to a theoretical dataset

from a matching protein digestion protocol, and a high score assigned by the database search engine indicates a strong match.

If no definitive identification can be made with a PMF, further peptide information can be obtain by analysing digested material using tandem mass spectrometry (MS/MS), which involves two stages of mass spectrometry. During the first stage, ionised intact peptides (often referred to as 'precursor ions') are measured before being isolated and fragmented. The fragmentation stage in MS/MS is most often via collision induced dissociation (CID), where the precursor ions collide with the molecules of an inert gas. The resulting fragments (known as the 'product ions') are measured in the second MS analyser, creating a peptide fragmentation pattern which can be searched against a database in a similar manner to a PMF search. The strength of a peptide match is based on the incidences of the observed precursor ion and product ion masses against the theoretical masses derived from the peptide sequence (Perkins *et al.* 1999). Again, a high score assigned by the database search engine indicates a strong match.

The identification of proteins significant in chemical signalling proves more difficult if there is little or no genomic data available for the species of interest, as there is limited database information to identify peptides and proteins against. In this case, MS/MS fragmentation spectra can be sequenced *de novo*, either manually or using software, where the mass difference between each product ion fragment denotes the mass of an amino acid. An example of the *de novo* sequencing of a peptide from an MS/MS spectrum is shown in Figure 1.6. The assembled amino acid sequence for that peptide can then be searched using the basic local alignment search tool for protein sequences (BLASTp), which searches against sequences in a selected database in the aim of identifying similar sequences or sequence tags that may denote the family membership of the protein (Atschul *et al*. 1990).

1.7 Mass spectrometry for protein quantification

The quantification of the signalling proteins in scent marks is essential since the regulation of the expression of these proteins is thought to be altered by a number of situations, including social setting and reproductive status (Stopka *et al*. 2007). Over the past decade, protein quantification methods have developed rapidly. Although more 'classical' quantification methods are still in use, such as western blotting and 2D gel analysis, mass spectrometry-based techniques have gained increased popularity for the analysis of proteins; due to the fact that these techniques are able

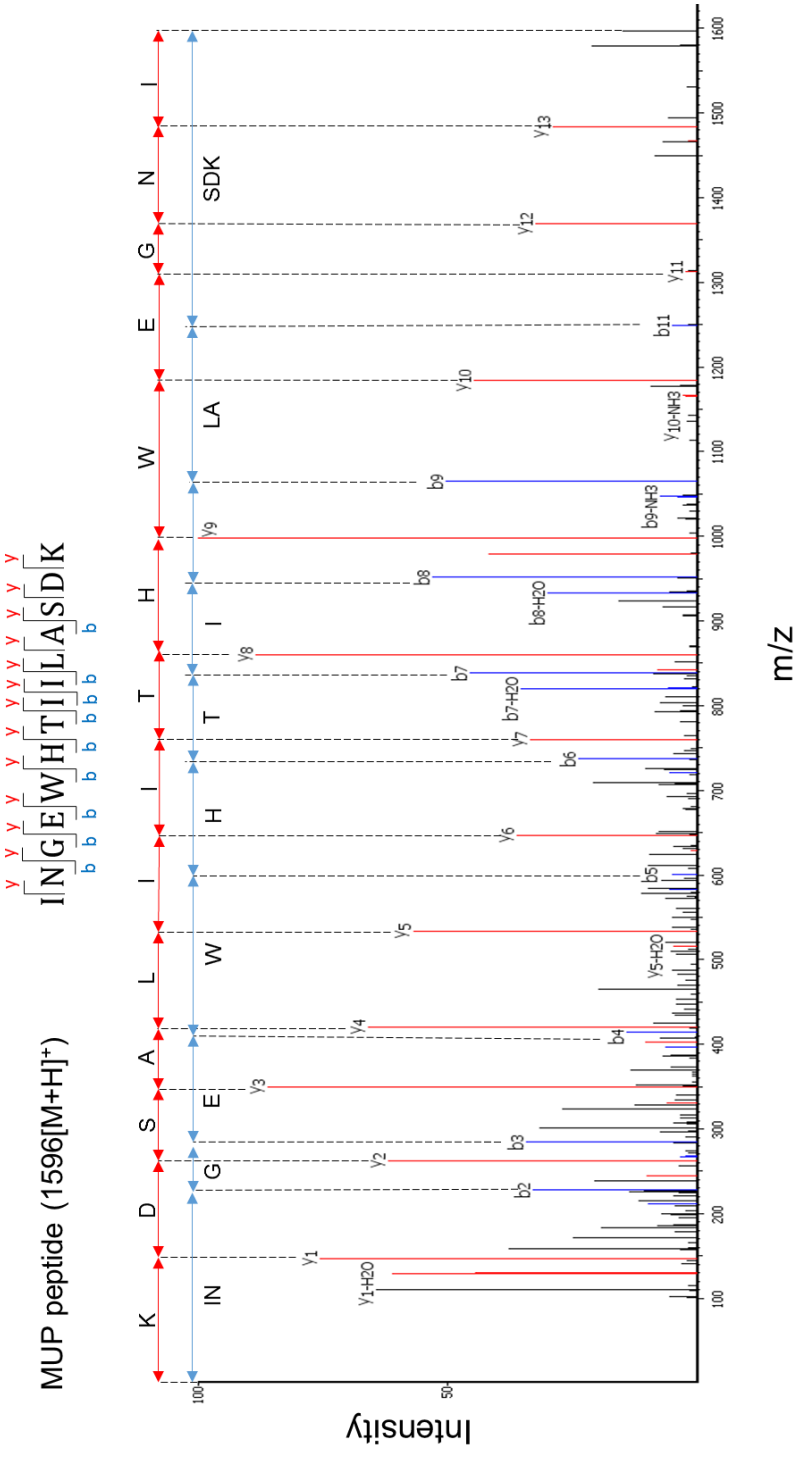


Figure 1.6 Example of *de novo* sequencing from MS/MS spectra.
In this example, the MS/MS spectrum has been used to assign a sequence to a known MUP peptide generated from a Lys-C digest. The mass difference between each y-ion is calculated and the amino acid residue corresponding to this mass difference is assigned. In this case, the y1-ion is at 147 Da since the peptide sequence is lysine terminated. The sequencing ends at the y14-ion at 1596 Da, which is the mass of the precursor ion. The assigned amino acid residues are read 'backwards' from the precursor ion mass for the correct sequence. The less intense b-ions can be used to confirm the sequence in the same way as the y-ions, except the sequence is to be read 'forwards' from the b1-ion. The MS/MS spectrum with y- and b-ion annotation, sequence determination and score indicating confidence that the correct sequence has been assigned was done using PEAKS software in this instance.

to not only quantify, but to also identify large groups of proteins (Bantschaff *et al*. 2007).

For mass spectrometry based quantification of proteins larger than around 15 kDa, such as MUPs, a 'bottom up' workflow is usually required, as the distribution of protein charge states renders 'top down' methods insensitive. 'Bottom up' protein analysis involves the proteolytic digestion of proteins prior to mass spectrometric analysis, whilst 'top down' involves isolating a protein ion for mass measurement and tandem mass spectrometry analysis (Brun *et al*. 2009). The quantitative data obtained can either be the absolute amount of protein in a sample, or the relative change of the amount of protein between two states. The data can be obtained via various methods, which can be label-mediated or label-free. The choice of quantification method to be used must be carefully considered prior to sample analysis.

1.7.1 Label-free quantification

In mass spectrometry based label-free quantitative proteomics, each sample of interest is prepared separately and analysed individually by LC-MS (Zhu *et al.* 2010). After fragmentation in the mass spectrometer, an ion with a particular mass-to-charge ratio (m/z) is detected, with the time that the ion arrives at the detector and its intensity being recorded. Label-free quantification is typically based on either peptide peak intensities in chromatography or the spectral counting of proteins post mass spectrometric analysis (Old *et al.* 2005; Silva *et al*. 2006; Zhu *et al.* 2010).

Relative label-free quantification of proteins is achieved via the comparison of the peak height of each peptide ion in multiple LC-MS datasets, as the signal intensity from electrospray ionisation correlates with ion (and therefore protein) concentration. The three most abundant peptides from each protein are used to calculate the relative amount of the protein (Zhu *et al.* 2010). Relative quantification of proteins can also be obtained by the comparison of the number of MS/MS spectra from a particular protein across numerous datasets. This is known as spectral counting, where the number of identified MS/MS spectra for a protein is determined by the number of unique peptides (Liu *et al*. 2004; Zhu *et al.* 2010). This results from the amount of coverage of a protein sequence, caused by the number of proteolytic peptides and therefore protein abundance. Comparing two sample types can therefore allow the relative amount of protein to be calculated, once normalisation and statistical analysis of the datasets has taken place (Zhu *et al.* 2010). These methods offer a simpler sample preparation workflow, which must be conducted extremely carefully to avoid errors.

A modified spectral counting strategy is used for the **absolute** quantification of proteins in label-free quantitative proteomics, termed absolute protein expression profiling, or APEX (Zhu *et al.* 2010). The absolute abundance of a protein per cell is calculated from the proportionality between the number of peptides observed and protein abundance; protein abundance being indicated by an APEX score calculated from the peptide mass spectra. Another method used for the absolute quantification of proteins is emPAI – the exponential form of PAI (protein abundance index). PAI is the number of identified peptides divided by the number of theoretical peptides for each protein, enabling approximate absolute protein abundance to be determined without undertaking any additional experimentation after a protein identification experiment (Rappsilber *et al*. 2002; Ishihama *et al*. 2005; Zhu *et al.* 2010). These absolute quantification methods suffer the same limitations as the previous relative quantification methods, but also allow the analysis of data from numerous experiments rather than samples analysed in a single experiment (Zhu *et al.* 2010).

1.7.2 Label-mediated quantification

The gold standard approach to quantitative proteomics is based on stable isotope dilution, which states that a peptide and its isotope-labelled analogue will behave in the same way when subjected to chromatographic and mass spectrometric analysis, due to their identical chemical nature (Brun *et al*. 2009). A known amount of the isotope-labelled peptide is added to the sample prior to analysis. In mass spectrometry, accurate peak ratios between the labelled and unlabelled peptides are observed, with the mass spectrometer being able to differentiate between their masses. Isotopic label based protein quantification exhibits good precision and linearity, however the accuracy depends on the method chosen, and all methods require complex sample preparation (Brun *et al*. 2009).

The **relative** quantification of proteins in label-mediated approaches involves calculating the ratio between the labelled and unlabelled peptides. Commonly used methods for relative quantification are isotope coded affinity tagging (ICAT) involving the labelling of cysteine residues in a sample using heavy and light isotope ICAT reagents to compare the relevant amounts of two samples (Gygi *et al*. 1999), and iTRAQ, which involves tagging amino acid N terminal and side chain amino groups, enabling 8 different samples to be quantified simultaneously due to the development of 8 different iTRAQ reagents (Boehm *et al*. 2007). Another method is stable isotope labelling by amino acids in cell culture (SILAC), a metabolic labelling technique where the labelled amino acids are grown in a medium with the cells of interest before being

pooled with unlabelled cells before being subjected to analysis (Ong *et al*. 2002). One cell culture is labelled with, for example, an amino acid in which every carbon atom is carbon-13 (heavy) and a second culture is labelled with the same amino acid in which every carbon atom is carbon-12 (light) (Claydon *et al.* 2011). This means that every peptide that contains that amino acid will appear as a heavy-light doublet in mass spectrometric analysis, and the relative intensities of these labelled and unlabelled ion peaks will indicate the relative expression of the protein (Claydon *et al.* 2011).

Absolute quantification methodologies based on stable isotope dilution strategies (Barr *et al*. 1996; Stocklin *et al*. 1997) rely on the sample of interest being spiked with known amounts of labelled analogues of specific peptides or proteins, which can be added before or after proteolysis (Brun *et al*. 2009).

A common method which involves adding the standards after proteolysis is the use of AQUA peptides. AQUA peptides are synthetic isotope-labelled peptides that correspond to the proteotypic peptides of interest, and are added to the sample just before LC-MS analysis (Brun *et al*. 2009). There is a limited choice of peptides that may be used as standards for quantification – peptides shorter than 15 amino acids in length are usually required, whilst reactive residues, such as methionine, should be avoided (Brun *et al*. 2009). For this reason, along with the cost of producing and quantifying each peptide, proteins are often quantified using a single AQUA standard, which is successful providing that complete proteolysis of the target protein has been achieved.

A method where the labelled standard is added prior to sample digestion involves the synthesis of an artificial concatamer of numerous different peptides of various protein targets. These standards are named QconCATs, and when digested release the isotope-labelled ‘Qpeptides’ which act as standards for the target peptides in the sample in LC-MS analysis (Beynon *et al*. 2005; Pratt *et al*. 2006; Brun *et al.* 2009). Quantification of analyte peptides by mass spectrometry is by the ratio between the analyte and QconCAT peptides (Pratt *et al*. 2006). This method is useful if analysing a set of related proteins, as analysis can take place in a single mass spectrometric run, reducing the potential for instrument error. However, suitable peptides must be chosen when designing a QconCAT, for example, peptides containing PTM sites or those that ionise poorly may not be suitable.

Finally, an increasingly popular quantification method utilises isotope-labelled full length proteins as standards for absolute quantification, termed PSAQ standards. PSAQ standards share the same biochemical properties as their targets, and so

behave identically to the proteins in the sample throughout sample preparation and MS analysis (Dupuis *et al*. 2008; Brun *et al*. 2009). In addition, these standards are spiked in defined amounts at the beginning of sample preparation; an advantage of this being the lack of differences between standard and sample digestion efficiency, providing more accurate quantification at MS level (Kaiser *et al*. 2011). Unlike AQUA standards, PSAQ standards have been shown to be compatible with any type of sample prefractionation method, including SDS-PAGE. Along with these various advantages over other absolute quantification methods, PSAQ provides greater sequence coverage due to the use of a full length protein standard, meaning isoforms and variants may also be distinguishable (Brun *et al*. 2009). Despite its numerous advantages, it must be remembered that PSAQ standards are currently costly and challenging to produce (Brun *et al*. 2009).

1.7 Aims

The project is based on the use of proteomic techniques to study communication in mice, involving the analysis of major urinary proteins (MUPs), whose function is chemical signalling amongst various *Mus* species. Mass spectrometry–based proteomics shall be used for three aspects of research into chemical signalling:

1. **For the discovery of MUPs in *Mus* species where no genome sequence is available.**

   Hypothesis: Mus spicilegus *mouse urine will contain MUPs not previously observed in other mouse populations.*

   The urinary proteins of *Mus spicilegus* are currently uncharacterised – the identification, and subsequent primary sequence determination, of major urinary proteins using various mass spectrometric methods may provide insight into their protein expression patterns, their roles in scent communication and their reproductive and social behaviour.

2. **To use stable isotope labelling of amino acids to track complex communication issues and investments.**

   Hypothesis: Whole animal dynamic labelling studies can be used to determine the proportion of lactation investment from communally nursing female house mice to ascertain any discriminatory factors through analysis of MUPs.

   In previous literature, it appears that communal nursing is more beneficial amongst related females due to a more mutually cooperative nursing system. It

could be that females recognise their own young and invest in those more than their non-offspring, and it could also be that females invest more in the offspring of their closer relatives (*i.e*. sisters) than their more distant relatives (cousins). It may be that females do not discriminate between their offspring and their non-offspring, and invest cooperatively more in larger, healthier pups. To answer these questions, whole animal metabolic labelling with stable isotope amino acids, introduced using a semi-synthetic diet, combined with the analysis of MUPs in female urine and various pup samples can provide relative proportions of investment from each mother to each pup. Heavy and light isotopes of the same amino acid are required, as this causes the resultant heavy/light peptides from the proteins expressed by the animal to behave the same under mass spectrometric analysis, yet provide a reasonable mass difference to differentiate between isotopes for calculation of label incorporation.

3. **To be able to quantify the dynamics of communication by absolutely quantifying MUPs – MUPs are extremely homologous, meaning quantitative approaches using MUP peptides is challenging.**

   Hypothesis: Intact mass is an appropriate method for the quantification of MUPs.

   Whilst intact mass analysis is an established valuable method for the identification of MUPs, the quantification of MUPs in a sample proves far more challenging – the high homology of MUP sequences has presented problems when implementing label-mediated absolute quantification methodologies at the peptide level. For protein quantification via electrospray ionisation of intact proteins, it would be essential to firstly determine whether the response factor varies for each protein, relating the signal intensity and ionisation efficiency observed to the concentration of the protein. Then, synthesis of the standard isotope labelled protein would be required – using a labelling method such as PSAQ. A labelled full–length protein would be particularly useful as a MUP quantification standard due to MUP sequence similarity. Labelling MUPs of a known concentration, which would then be spiked into an unknown MUP containing solution, would possess the same chemical properties as their unlabelled analogues, but a notably different mass, enabling a difference in mass to be seen in a mass spectrum. A number of synthetic MUPs of known concentration could be made and used as internal standards for electrospray ionisation of intact proteins.

Chapter 2: Methods

2.1 Sample collection

2.1.1 *Mus spicilegus* mice

Urine was collected using a recovery method. Mice were individually placed on a wire grid over a clear plastic container, with another over the top to confine the animal. Urine was collected by pipetting from the container and samples were stored at -20 °C in microcentrifuge tubes until required. Mice were housed in a temperature (20 °C) and humidity controlled (50 – 60 °C) environment with a 12 hour light cycle (12 hours of light/12 hours of darkness). Males were housed individually; females were housed in groups of two to three per cage.

2.1.2 C57BL/6 and BALB/c laboratory mice

Urine was collected from adult C57BL/6 and BALB/c mice by gentle bladder massage by members of staff based at Leahurst campus, University of Liverpool. Urine was collected in microcentrifuge tubes and stored at -20 °C until required for analysis. Mice were housed in a temperature (20 °C) and humidity controlled (50 – 60 °C) environment with a 12 hour light cycle (12 hours of light/12 hours of darkness). BALB/c pup samples (stomach contents, urine, liver and muscle) were recovered from pups by members of staff based at Leahurst campus, University of Liverpool, and tissue samples were rinsed with phosphate buffered saline, pH 7.5 to remove the majority of residual blood. These samples were stored in microcentrifuge tubes at -20 °C until required for analysis.

2.2 Preparation of labelled diet for BALB/c laboratory mice

Labelled diets were prepared by Dr J. P. Green of the Mammalian Behaviour and Evolution group, University of Liverpool. 5002 LabDiet® (Purina Mills PMI®), in pellet form, was used to separately prepare [$^{2}H_{8}$] valine, [$^{13}C_{6}$] lysine, [$^{2}H_{4}$] lysine and [$^{2}H_{9}$] lysine labelled diets. The amount of valine or lysine present in 1 kg of diet was determined, and the equivalent amount of the required labelled amino acid was dissolved in reverse osmosis (RO) purified water (1 L). 1 kg of 5002 LabDiet® was stirred into the mixture and left for two hours. A further 125 ml RO water was added to the mixture, which was then stirred before being left for a further hour. The mixture was then transferred to a blender along with 125 ml RO water, and blended until smooth. A piping bag was used to pipe the labelled diet onto dehydrator trays lined

with baking paper, and pellet lines were scored across the diet. The diet was placed in a dehydrator for approximately 48 hours at 40 °C before being split into pellets.

2.3 Protein assay of mouse urine and tissue samples

Total protein concentration in mouse urine and tissue samples was measured using a Coomassie Plus protein assay kit (Pierce, Rockford, USA). A standard curve, ranging from 0-50 µg/ml, was produced by diluting a stock solution of Bovine Serum Albumin (BSA, 1 mg/ml) with MilliQ water. Samples were diluted as required with MilliQ water. Absorbance readings were measured at 620 nm using a Thermo Scientific™ Multiskan™ plate reader, and sample dilutions were accounted for when calculating total protein amounts in samples.

2.4 Creatinine assay of mouse urine

Creatinine in mouse urine was quantified using a creatinine assay kit from Sigma, UK. The creatinine standard curve ranged from 0-30 µg/ml. Urine samples were diluted with MilliQ water as required, and absorbance readings were measured at 570 nm using a Thermo Scientific™ Multiskan™ plate reader. Sample dilutions were accounted for when calculating total creatinine amounts in urine.

2.5 SDS-PAGE

SDS-PAGE was performed based on the method described by Laemmli (1970). Approximately 10 µg of protein was mixed 1:1 with reducing sample buffer and loaded onto a 15% (w/v) acrylamide gel. Electrophoresis was carried out at a constant voltage of 200 V. The SDS gels were stained, allowing protein bands to be visualised, using Coomassie blue stain (Sigma) for several hours before being destained in 80:10:10 MilliQ water:methanol:acetic acid.

2.6 Urine fractionation by strong anion exchange chromatography

Separation of individual MUPs from *Mus spicilegus* urine and C57BL/6 laboratory mouse urine was by strong anion-exchange chromatography with UV detection using a Dionex BioLC system. Pooled urine (typically 15-20 µl) was desalted using Thermo Zeba spin columns (7 kDa MWCO) and loaded onto a Thermo Dionex Propac™ SAX-10 2×450 mm column using a mobile phase of 25 mM Tris HCl pH 8.8. Elution took place over a 25 minute gradient from 0%-30% 25 mM Tris HCl buffer, pH 8.8 with 1 M NaCl using a flow rate of 0.2 ml/min. Eluent from the column was monitored at 280 nm in a 9 mm path length flow cell. Fractions were manually collected in microcentrifuge tubes and run on a SDS-PAGE gel to determine the protein containing fractions.

2.7 ESI-MS of intact proteins

Mus spicilegus urine samples were diluted to approximately 2 pmol/µl in 0.1% (v/v) formic acid before analysis using a Waters nanoAqcuity UPLC coupled to a Waters Synapt™ G1 Q-ToF mass spectrometer fitted with an electrospray source (Waters, Manchester, UK). The mass spectrometer was externally calibrated with horse heart myoglobin (1 pmol/µl, Sigma). Diluted urine was injected onto a Waters MassPREP™ micro desalting column (2.1 x 5 mm, 20 µm particle size, 1000 Å pore size) and protein was eluted using a mixture of two solvents - solvent A (HPLC grade water with 0.1% (v/v) formic acid) and solvent B (HPLC grade acetonitrile with 0.1% (v/v) formic acid). Separations were performed using a stepwise gradient of 5% to 95% solvent B over 10 min at a flow rate of 40 µl/min, and mass spectra were acquired between m/z 300 – 2000. Mass spectra were processed, deconvoluted and protein molecular weight calculated using Waters MassLynx 4.1 software.

C57BL/6 urine samples and recombinant MUP samples were diluted as required in 0.1% (v/v) formic acid before analysis using a Waters nanoAqcuity UPLC coupled to a Waters Synapt™ G1 or G2 Q-ToF mass spectrometer fitted with an electrospray source (Waters, Manchester, UK). The mass spectrometer was externally calibrated with horse heart myoglobin (500 fmol/µl, Sigma). Samples were injected onto a Waters MassPREP™ micro desalting column (2.1 x 5 mm, 20 µm particle size, 1000 Å pore size) and protein was eluted using a mixture of two solvents - solvent A (HPLC grade water with 0.1% (v/v) formic acid) and solvent B (HPLC grade acetonitrile with 0.1% (v/v) formic acid). Separations were performed using a stepwise gradient of 5% to 95% solvent B over 12 min at a flow rate of 20 µl/min, and mass spectra were acquired between m/z 300 – 2000. Mass spectra were processed, deconvoluted and protein molecular weight calculated using Waters MassLynx 4.1 software.

2.8 Homogenisation of mouse tissue samples

Liver and muscle samples from BALB/c pups were weighed and placed into microcentrifuge tubes, and kept on ice. 9 volumes of 25 mM NH_4HCO_3 (or 18 volumes if tissue sample weighed less than 20 mg) (4 °C) was added to each tissue sample and samples were homogenised manually in the microcentrifuge tube using a small plastic pestle. The pestle was rinsed with 25 mM NH_4HCO_3 between each sample. Homogenates were centrifuged at 16,000 x *g* for 15 minutes at 4 °C, and all supernatant was transferred to a new microcentrifuge tube before being stored at -20 °C until required for analysis.

2.9 Protein digestion

2.9.1 In-gel protein digestion

Pieces of protein bands (plugs) from SDS-PAGE analysis were excised and placed in microcentrifuge tubes, then destained in 50:50 ACN:25 mM NH_4HCO_3 for 15 minutes at 37°C. This was repeated until all stain was removed from the plugs. Plugs were reduced in 25 µl of 10 mM dithiothreitol (DTT) at 60 °C for 1 hour, to reduce the disulfide bonds between the cysteine residues in the protein. The DTT was discarded and 25 µl of 55 mM iodoacetamide was added to the plugs to prevent the re-formation of the disulfide bonds between cysteine residues. This alkylation step was carried out in the dark at room temperature for 45 minutes. The gel plugs were then dehydrated in acetonitrile for 15 minutes at 37 °C. 10 µl of 10 ng/ml endoprotease Lys-C was added to each of the gel plugs and samples were incubated at 37 °C for 16 hours. The digestion reaction was stopped with the addition of formic acid (1% v/v).

2.9.2 In-solution protein digestion of mouse urine and SAX chromatography fractions

Urine samples were diluted in 25 mM NH_4HCO_3 to a concentration of 10 µg/µl of protein. Samples were incubated for 10 minutes with 0.1% w/v *Rapi*Gest™ SF Surfactant (Waters) at 80 °C. The samples were then reduced for 10 minutes with DTT (3 mM final concentration) at 60 °C, followed by alkylation with iodoacetamide (9 mM final concentration) in the dark at room temperature for 30 minutes. 2.5 µl of protease (endoprotease Lys-C or endoprotease Glu-C, 0.1 µg/µl diluted in 25 mM Tris HCl pH 8.5), was added to the samples which were then incubated at 37 °C for 16 hours. Following incubation, 5 µl of each digested sample was taken to run on an SDS-PAGE gel to check for complete digestion. The remaining digest was treated with trifluoroacetic acid, final concentration of 0.5% (v/v) and incubated at 37 °C for 45 minutes to hydrolyse the *Rapi*Gest™ SF Surfactant prior to LC-MS analysis. Samples were centrifuged at 7,000 x *g* for 15 minutes and the supernatant transferred to a new microcentrifuge tube.

2.9.3 In-solution protein digestion of homogenised mouse tissue samples

For each sample, approximately 100 µg of protein (as determined by protein assay, section 2.3) was made up to a volume of 160 µl with 25 mM NH_4HCO_3. Samples were incubated for 10 minutes with 1% w/v *Rapi*Gest™ SF Surfactant (Waters) at 80 °C. The samples were then reduced for 10 minutes with DTT (3 mM final

concentration) at 60 °C, followed by alkylation with iodoacetamide (9 mM final concentration) in the dark at room temperature for 30 minutes. 2.5 µl of endoprotease Lys-C (0.1 µg/µl diluted in 25mM Tris HCl pH 8.5), was added to the samples which were then incubated at 37 °C for 16 hours. Following incubation, 5 µl of each digested sample was taken to run on an SDS-PAGE gel to check for complete digestion. The remaining digest was treated with trifluoroacetic acid (to a final concentration of 0.5% (v/v)) and incubated at 37 °C for 45 minutes to hydrolyse the *Rapi*Gest™ SF Surfactant prior to LC-MS analysis. Samples were centrifuged at 7,000 x *g* for 15 minutes and the supernatant transferred to a new microcentrifuge tube.

2.10 Peptide mass fingerprinting by MALDI-ToF-MS

Peptide mixtures from in-gel and in-solution digestion of proteins were mixed 1:1 with a MALDI matrix (10 mg/ml *α*-cyano-4-hydroxycinnamic acid in 50% ACN/0.1% TFA) and were spotted onto a target plate which was left to air dry. MALDI-ToF-MS was carried out using a Bruker UltrafleXtreme™, which was operated in reflectron mode with positive ion detection. The mass spectrometer was externally calibrated with a mixture of stndards from Sigma - Des-Arg bradykinin (904.47 Da), angiotensin I (1296.69 Da), neurotensin (1672.92 Da), ACTH 1-17 fragment (2093.09 Da) ACTH (corticotrophin, 2465.2 Da) and ACTH 7- 38 fragment (3657.93 Da). All standards apart from ACTH 7 – 38 (at 9 pmol/µl) were at a concentration of 6 pmol/µl. Spectra were acquired between 800 – 4000 m/z. Laser frequency was set to 1000 Hz and laser energy was set between 27 – 33 % of the maximum energy, with 500 shots per spectrum. Spectra were then searched against an appropriate database via the Mascot search engine to identify peptides and produce a peptide mass fingerprint.

2.11 Protein discovery by LC-MS/MS

Digested protein samples were diluted to approximately 500 fmol/µl in 97:3 water:ACN + 0.1% (v/v) formic acid before analysis using a Waters nanoAqcuity UPLC coupled to a Waters Synapt™ G2 Q-ToF mass spectrometer fitted with a nanospray source (Waters, Manchester, UK). The mass spectrometer was operated in positive ion MS^E mode, with the conditions for analysis set as follows: capillary voltage - 3 kV; cone voltage - 45 V; source temperature - 80 °C; desolvation temperature - 150 °C; cone gas flow - 50 L/hr; desolvation gas flow - 500 L/hr. The mass spectrometer detectors were calibrated with Leucine-enkephelin (50 pmol/µl) (Waters, Manchester, UK), and Glu-fibrinopeptide (5 pmol/µl) (Waters, Manchester, UK) was used for the mass calibration. Samples were injected onto a Waters C18 trapping column (180 µm x 20 mm) before being separated using an ACQUITY

UPLC® BEH column C18 analytical column (75µm x 150mm, 1.7µm) and a mixture of two solvents - solvent A (HPLC grade water with 0.1% (v/v) formic acid) and solvent B (HPLC grade acetonitrile with 0.1% (v/v) formic acid). Separations were performed using a linear gradient of 3% to 85% solvent B over 35 minutes at a flow rate of 300 nl/min, and mass spectra were acquired between m/z 300 – 3000. Mass spectra were viewed using Waters MassLynx 4.1 software.

2.12 Discovery and *de novo* sequencing by LC-MS/MS

Digested protein samples were diluted to approximately 500 fmol/µl in 97:3 water:ACN + 0.1% (v/v) formic acid before analysis using an Ultimate 3000 nano system (Dionex/Thermo Fisher Scientific, Hemel Hempstead, UK) coupled to a QExactive mass spectrometer (Thermo Fisher Scientific, Hemel Hempstead, UK). The mass spectrometer was calibrated using a positive ion calibration solution containing a mixture of caffeine, MRFA, Ultramark1621 and n-butylamine (Thermo Fisher Scientific, Hemel Hempstead, UK). The mass spectrometer was operated in data dependent ESI+ mode, which automatically switched between MS and MS/MS acquisition. Survey full scan MS spectra (300-2000 m/z) were acquired in the Orbitrap with 70,000 resolution (200 m/z) and dynamic exclusion was set to 20 seconds. The 10 most intense multiply charged ions ($z \geq 2$) were sequentially isolated and fragmented in the octapole collision cell by higher energy collisional dissociation (HCD) with a fixed injection time of 120 milliseconds and a resolution of 35,000. Conditions for the analysis were as follows: spray voltage - 1.9 kV (no sheath or auxillary gas flow); heated capillary temperature – 250 °C; normalised HCD collision energy - 30%. The MS/MS ion selection threshold was set to 1 x 10^4 ion counts and a 2 *m/z* isolation width was set. Samples were injected onto a C18 Acclaim PepMap 100 trap columnn (2 cm x 75 µm inner diameter, 3 µm, 100Å) (Dionex) before being separated using a C18 Easy-Spray PepMap® RSLC (15 cm x 75 µm inner diameter, 2 µm, 100Å) (Dionex) and a mixture of two solvents - solvent A (HPLC grade water with 0.1% (v/v) formic acid) and solvent B (HPLC grade acetonitrile 80% (v/v) with 0.1% (v/v) formic acid). Separations were performed using a linear gradient of 3.8% to 50% solvent B over 35 min at a flow rate of 300 nl/min.

2.13 *De novo* sequencing analysis using PEAKS 6

De novo sequencing analysis of mass spectrometric data from digested *Mus spicilegus* urine and digested fractions from SAX chromatography of *Mus spicilegus* urine (acquired using the QExactive) was assisted using PEAKS 6 software

(Bioinformatics Solutions Inc, Canada). Data were processed using a precursor error tolerance of 10 ppm, and a fragment ion error tolerance of 0.01 Da. Post translational modifications (carbamidomethylation – fixed, oxidation of methionine residues – variable) and fragmentation type of high-energy C-type dissociation (HCD) were also included in the processing parameters. The confidence score assigned by the software which indicates the likelihood of a peptide being assigned the correct sequence was set at a minimum 'score' of 55%, which is the value recommended by PEAKS.

2.14 Bacterial transformation of rMUPs

The transformation process was carried out by Mrs L McLean, University of Liverpool. The genes encoding MUPs 4, 7, 11 and 20 (darcin) were cloned into *E. coli* expression vectors pET28b resulting in the first fourteen residues of the recombinant proteins corresponding to the cloning and N-terminal hexa-histidine purification tag which are not part of the native sequences. Transformation was into *E. coli* BL21 (DE3) cells at 37 °C. Competent *E. coli* BL21 (DE3) cells were thawed on ice for 10 minutes before being mixed, then 50 µl of each were transferred to separate tubes and kept on ice. Plasmid DNA (5 µl) was added to each cell mixture, the contents were mixed, and were placed on ice for 10 minutes. The mixtures were heat shocked at 42 °C, using a water bath, for 10 seconds, then placed on ice for a further 5 minutes. A super optimal broth (SOC) solution (950 µl), containing 2% (w/v) tryptone, 0.5% (w/v) Yeast extract, NaCl (10 mM), KCl (2.5 mM), $MgCl_2$ (anhydrous 10 mM) and deionised water, was added to the mixtures. The SOC solution was supplied by Promega UK. The mixtures were incubated at 37 °C for 60 minutes, on a mixer at 250 rpm. Cells were then mixed by inversion and diluted ten-fold in SOC. LB agar plates were heated to 37 °C before the diluted transformation mixtures (50 µl) were added to the plates, which were then incubated at 37 °C overnight. The next day, glycerol stocks of the plasmids were produced for long-term storage. A single colony from each LB plate was added to a culture of LB medium (5 ml) containing ampicillin (100 µg/ml). The cultures were incubated at 37 °C for 6 hours, shaking at 300 rpm. The bacterial cells were then harvested by centrifugation (15 minutes at 2400 x *g*) at 4 °C. The centrifuged bacteria were then added 1:1 to sterilised 60% glycerol solutions. The glycerol bacterial stocks were aliquoted (100 µl) and stored at -80 °C prior to protein expression.

2.15 Expression and purification of recombinant MUPs

Expression of recombinant MUPs 4, 7, 11 and 20 was carried out by Mrs L. McLean, University of Liverpool. Glycerol stocks were defrosted and streaked onto LB agar plates containing ampiciliin (50 mg/ml), using a loop and sterile technique. The plates were then incubated at 37 °C overnight. The next day, individual colonies were incubated in LB broth (10 ml) and ampicillin (10 µl, 50 mg/ml) at 37 °C for six hours. Each LB culture (100 µl) was added to minimal media containing disodium phosphate (0.24 M), potassium phosphate (0.11 M), sodium chloride (11 mM), ammonium chloride (93 mM), magnesium sulphate (1 M), calcium chloride (0.1 M), glucose (20%, 1 g in 5 ml), thiamine (0.5 % (w/v)) and deionised water. The cultures were incubated at 37 °C, shaking at 300 rpm, overnight. The next day, 6 ml of each culture was added to 200 ml of minimal media plus a full set of unlabelled amino acids (10 mg/ml of hydrophilic amino acids and 20 mg/ml of hydrophobic amino acids). The cultures were incubated at 37 °C, shaking at 300 rpm, for 40 minutes. The absorbance readings of the cultures (at 600 nm) were taken every hour until it absorbance readings reached 0.6, with the minimal media used as a reference. MUP expression was induced with isopropyl-D-thiogalactopyranoside (IPTG) and cells were harvested by centrifugation for 15 minutes (3500 rpm) at 4 °C. Inclusion bodies containing the proteins were recovered using BugBuster Protein Extraction Reagent (Novagen, Nottingham, UK). Inclusion bodies were re-suspended in 80 mM phosphate buffer, 6 M guanidinium chloride, 2 M NaCl, 40 mM imidazole, pH 7.4. Each recombinant MUP was then purified from these solutions using affinity chromatography with a nickel-based resin (HisTrap kit, GE Healthcare, Bucks., UK). Protein samples were loaded and then the HisTrap columns were washed with 80 mM phosphate:2 M NaCl buffer, pH 7.4. Protein samples were then eluted with 80 mM phosphate:2 M NaCl: 0.5 M imidazole buffer, pH 7.4 during which 1 ml fractions were collected. Purified rMUPs were then desalted by dialysis against 100 volumes of 100 mM ammonium bicarbonate buffer, pH 8.5, for three hours.

2.16 Database searching

2.16.1 PEAKS 6

Raw mass spectra from QExactive analyses were imported into PEAKS 6 software (Bioinformatics Solutions Inc, Canada). Peptides (and the proteins from which they were generated) were identified by searching against a Uniprot mouse database, a mature MUPs database and a custom-made *Mus spicilegus* MUPs database. The *Mus spicilegus* MUPs database was added to the PEAKS software using the

sequences constructed for the *Mus spicilegus* MUPs in FASTA format. Spectra were searched using a precursor error tolerance of 10 ppm, and a fragment ion error tolerance of 0.01 Da, and to allow 1 missed cleavage. Post translational modifications (carbamidomethylation – fixed, oxidation of methionine residues – variable) and fragmentation type of high-energy C-type dissociation (HCD) were also included in the processing parameters.

2.16.2 Protein Lynx Global Server (PLGS)

Raw mass spectra from Synapt G2 analyses were imported into Waters PLGS (version 2.5.2) software. Peptides (and the proteins from which they were generated) were identified by searching against a database of mouse proteins obtained from Uniprot. The parameters set for the database searches were an FDR of 4%, to allow 1 missed cleavage, a minimum of 1 peptide match per protein, a minimum of 3 fragment ion matches per peptide, and a minimum of 7 fragment ion matches per protein. Post translational modifications (carbamidomethylation – fixed, oxidation of methionine residues – variable) were also included in the processing parameters.

2.16.3 Mascot

Raw mass spectra from Bruker UltrafleXtreme™ analyses were imported into Biotools™ (Bruker) and then searched using Mascot (Matrix Science). Peptide mass fingerprint searches were carried out searching against the Swissprot database (taxonomy – *Mus musculus*). The parameters set for the database searches were to allow 1 missed cleavage, a peptide mass tolerance ± 100 ppm, a peptide charge state of 1+, and post translational modifications (carbamidomethylation – fixed, oxidation of methionine residues – variable).

2.17 Determination of label incorporation in adult female BALB/c urine

Raw mass spectra from Bruker UltrafleXtreme™ analyses were imported into Biotools™ (Bruker) and then searched using Mascot (Matrix Science). MUP peptides were identified by peptide mass fingerprint searches (as described in section 2.16.3). The corresponding labelled MUP peptides were identified from the 'light' (unlabelled) peptide masses, their sequence information, and the mass shift expected for the 'heavy' (labelled) peptide. The label incorporation into MUP peptides was determined by calculating the precursor relative isotopic abundance (RIA):

$$RIA = I_H / (I_H + I_L)$$

where I_H is the signal intensity displayed by the heavy labelled peptide (calculated by summing the intensities of the monoisotopic (M_0), M_0 + ^{13}C (M_1) and M_0 + 2 ^{13}C (M_2)

isotopic peaks), and I_L is the signal intensity of the light (unlabelled) peptide, calculated in the same way.

2.18 Determination of label incorporation in pup urine, stomach contents and tissue samples

Raw mass spectra from Synapt G2 analyses were imported into Waters PLGS (version 2.5.2) software. Unlabelled and their corresponding [$^{13}C_6$] lysine labelled peptide masses were identified by database searches as described in section 2.16.2, including [$^{13}C_6$] lysine as a post translational modification in the processing parameters. The label incorporation into peptides of interest was determined by calculating the precursor RIA, as described in section 2.17.

Chapter 3: Major Urinary Proteins in *Mus spicilegus*

3.1 Introduction

Mus spicilegus, also known as the mound-building mouse, is a non-commensal species from the steppe grassland habitats in eastern Europe (Bonhomme 1992, Sokolov *et al.* 1998; Patris *et al*. 2002). *M. spicilegus* are unique in their construction of mounds, which are thought to provide shelter and protection during the winter months, and are believed to be built and inhabited by members of the same family. Despite *M. spicilegus* being genetically close to the *Mus musculus* group of the *Mus* species group (Figure 3.1), the mating patterns and social behaviour observed in *M. spicilegus* are very different to those seen in the house mouse, *Mus musculus domesticus*. *M. spicilegus* female mice show a preference for one particular mate (suggestive of a possible monogamous mating system), whilst *M. m. domesticus* exhibits polygamous mating behaviour (Patris and Baudoin 1998; Patris *et al*. 2000).

Unlike *M. m. domesticus*, where females have a sexual preference for dominant males and have developed a communal nesting and nursing system, *M. spicilegus* are not cooperative breeders; they exhibit strong mating pair bonds and females appear to be very aggressive towards each other (Tong *et al*. 2012). In studies comparing behavioural characteristics of *M. m. domesticus* and *M. spicilegus*, many differences have been observed, in particular their tolerance towards siblings and strangers (Sczenczi *et al*. 2012). In the behavioural study conducted by Szenczi *et al*. (2012) (Figure 3.2), male *M. m. domesticus* expressed more time in agonistic behaviour towards both siblings and strangers to a similar extent. Male *M. spicilegus* showed an increase in time spent in agonistic behaviour with age, with a very significant difference in the time spent in offensive and defensive behaviour towards strangers compared to siblings. With increasing age, female *M. m domesticus* and female *M. spicilegus* expressed agonistic behaviour towards strangers, but none towards siblings. Female *M. spicilegus* display significantly more agnostic behaviour towards strangers than female *M. m. domesticus*. Also apparent is that female *M. spicilegus* display low levels of tolerance not only to unfamiliar females, but also to unfamiliar males. This, along with the higher levels of cooperation in terms of paternal care compared to *M. m. domesticus*, suggest the possibility of a monogamous mating system (Gouat *et al*. 2003).

Relatively little is known regarding the role of chemical communication in the unique

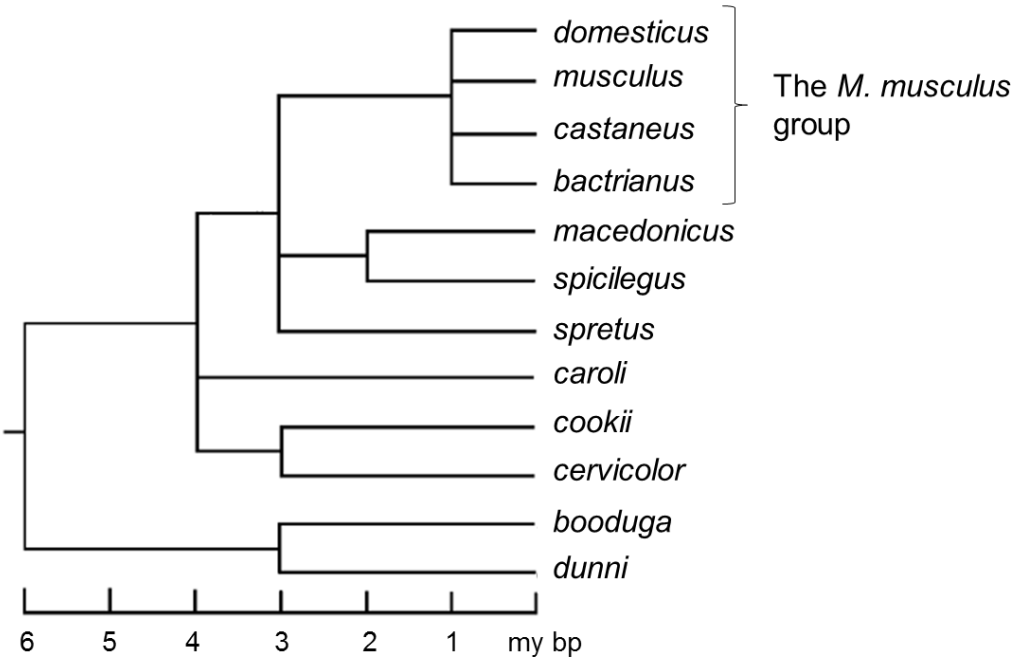


Figure 3.1 Phylogenetic tree of *Mus* species.
Domesticus, musculus, castaneus and *bactrianus* are all members of the *Mus musculus* subspecies group. *Spicilegus* are most closely related to *macedonicus*, but also closely related to *spretus* and the *Mus musculus* group. Estimated times of divergence between evolutionary lines are given as the number of millions of years before present (my bp). Adapted from Silver 1995.

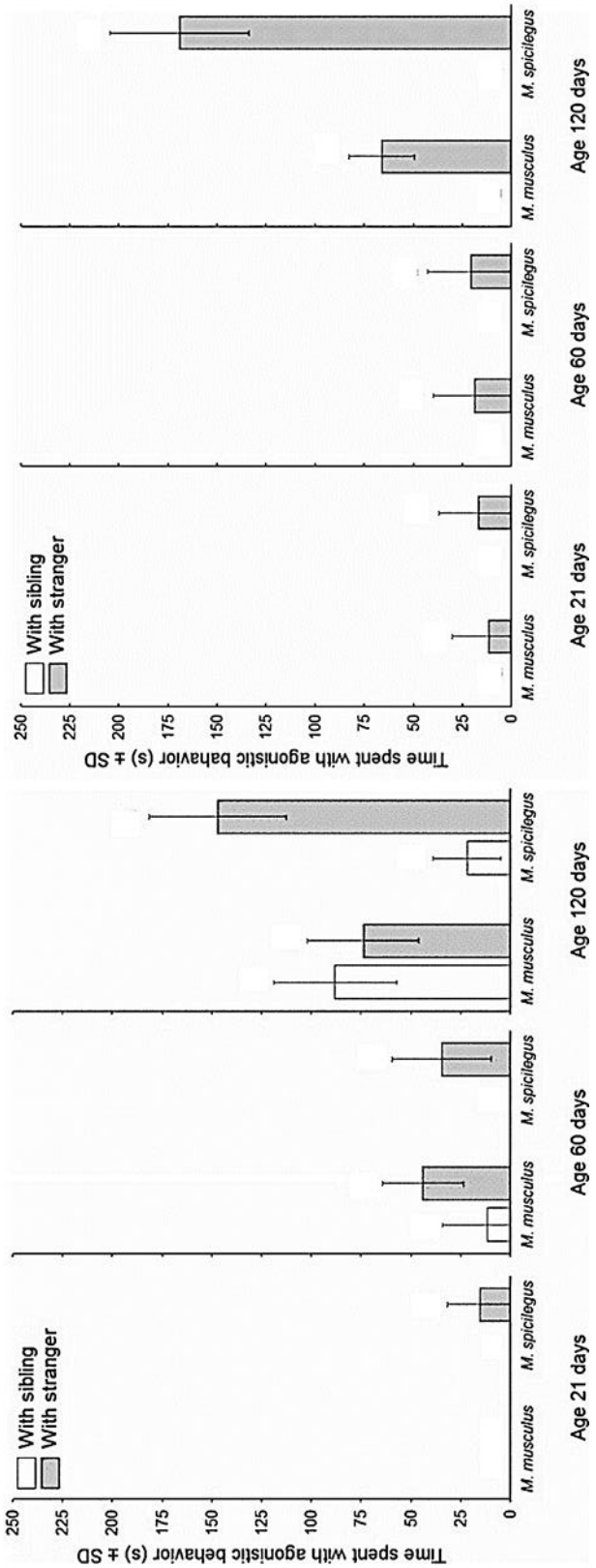


Figure 3.2 Taken from Szenczi *et al.* Average time spent in offensive and defensive behaviour by male *M. spicilegus* and *M. m. domesticus* (left) and by female *M. spicilegus* and *M. m. domesticus* (right). With increasing age, male *M. m. domesticus* express more time in agonistic behaviour towards both siblings and strangers to a similar extent. Male *M. spicilegus* also show an increase in time spent in agonistic behaviour with age, with a very significant difference in the time spent in offensive and defensive behaviour towards strangers compared to siblings. At 120 days old, male *M. spicilegus* display significantly more agonistic behaviour towards strangers than male *M. m. domesticus*, and significantly less to siblings. With increasing age, female *M. m. domesticus* and female *M. spicilegus* express more time in agonistic behaviour towards strangers, but none towards siblings. At 120 days old, female *M. spicilegus* display significantly more agnostic behaviour towards strangers than female *M. m. domesticus*.

social and reproductive behaviours displayed by *M. spicilegus* (Tong *et al*. 2012). As urinary MUPs play a significant part in the social and reproductive behaviours displayed by mice such as *M. m. domesticus*, this study sought to evaluate whether *M. spicilegus* also invested in MUPs, and by inference, whether these could play a key role in their unusual behaviours. To do this, the protein content of the urine of male and female *M. spicilegus* was examined, and then mass spectrometry was used to identify the proteins and characterise their primary structure. The possible roles of these proteins in scent communication was then investigated by observing changes in urinary protein expression in experiments relating to their sexual and social behaviour.

3.2 Results and discussion

3.2.1 The protein content of *Mus spicilegus* urine

Male mice were singly housed, and female mice were housed in sibling groups. Urine was collected from a number of males and females as described in Chapter 2, and was initially analysed by SDS-PAGE. An intensely staining group of proteins at approximately 20 kDa were identified in the male *M. spicilegus* urine, the number and intensity of which varied between individuals. A single, low intensity protein band of similar mobility was identified in all female *M. spicilegus* urine samples (Figure 3.3). Protein assays of the male and female urine samples confirmed varied protein concentrations in male urine (0.5 mg/ml – 2 mg/ml), and less varied in female urine (0.7 mg/ml – 1 mg/ml) (Figure 3.4).

3.2.2 Determination of accurate molecular weight of proteins of interest using ESI-MS

To obtain a more accurate molecular weight of the proteins observed in SDS-PAGE, desalted urine was analysed by ESI-MS (Figure 3.5). Three clearly resolved masses (18585 Da, 18742 Da and 18762 Da) were identified in the male *M. spicilegus* samples. These protein masses are between 18 – 19 kDa, as are the masses of known MUPs from other *Mus* species (Beynon *et al.* 2014). The intensity of these three masses varied between male samples, and in some individuals, a fourth mass at 18918 Da was present. None of the masses identified in *M. spicilegus* urine matched those of known *M. m. domesticus* MUPs. Several of the female *M. spicilegus* urine samples revealed a protein mass (18918 Da) that was also present in some male samples. However, due to the low concentration of protein in female urine, it was not possible to identify the protein in all samples.

From SDS-PAGE and intact protein analysis, protein expression patterns in *M. spicilegus* urine appear less complex than in *M. m. domesticus* urine (Robertson *et. al*. 2007) but there was greater variation in protein expression between the male *M. spicilegus* than between inbred *M. m. domesticus* individuals (Cheetham *et. al.* 2009).

3.2.3 Peptide mass fingerprinting of urine samples

Pieces of SDS-PAGE gel from the region containing the 20 kDa bands were subjected to in-gel digestion with endoprotease Lys-C. Endoprotease Lys-C was chosen as it results in fewer, larger peptides than trypsin, suitable for analysis using MALDI-TOF-MS (Robertson *et. al.* 2007). Comparison of the peptide mass fingerprints (PMFs)

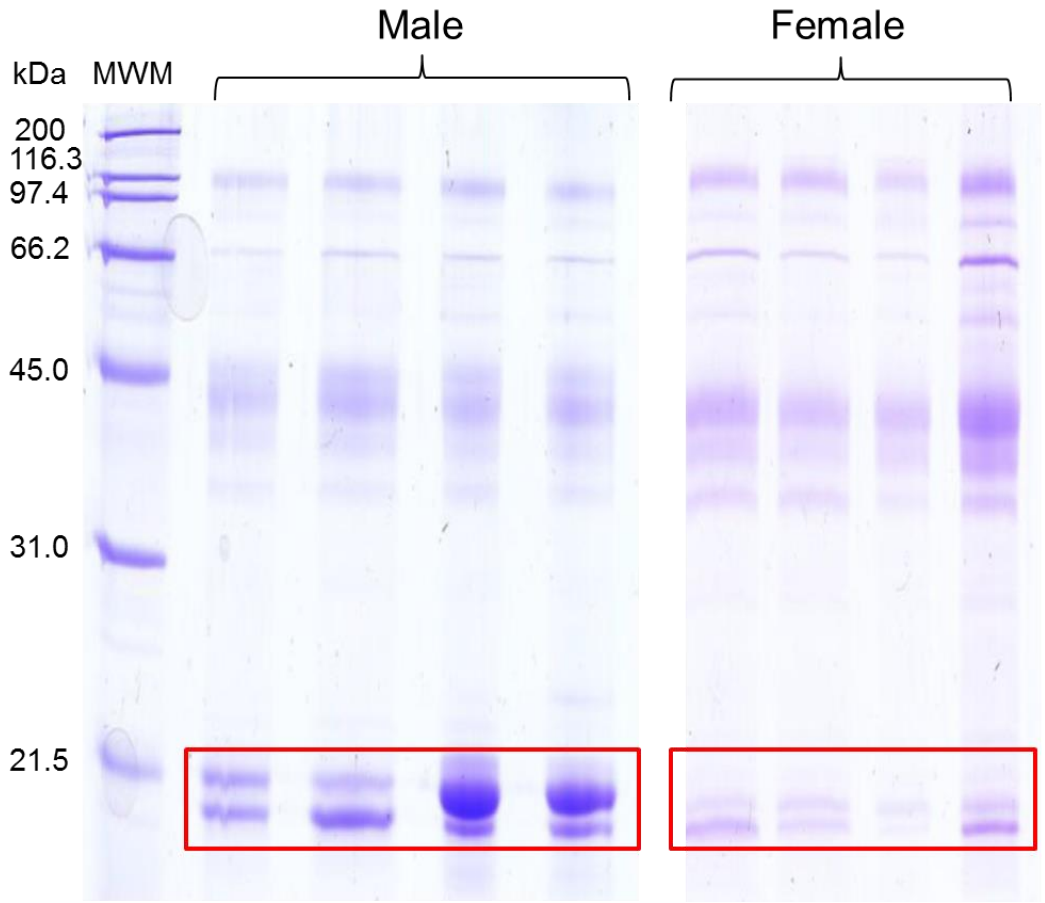


Figure 3.3 SDS-PAGE analysis of male and female *Mus spicilegus* urine.
Urine (5 µl) from both male and female *Mus spicilegus* was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. The protein bands of interest at approximately 20 kDa are highlighted in red.

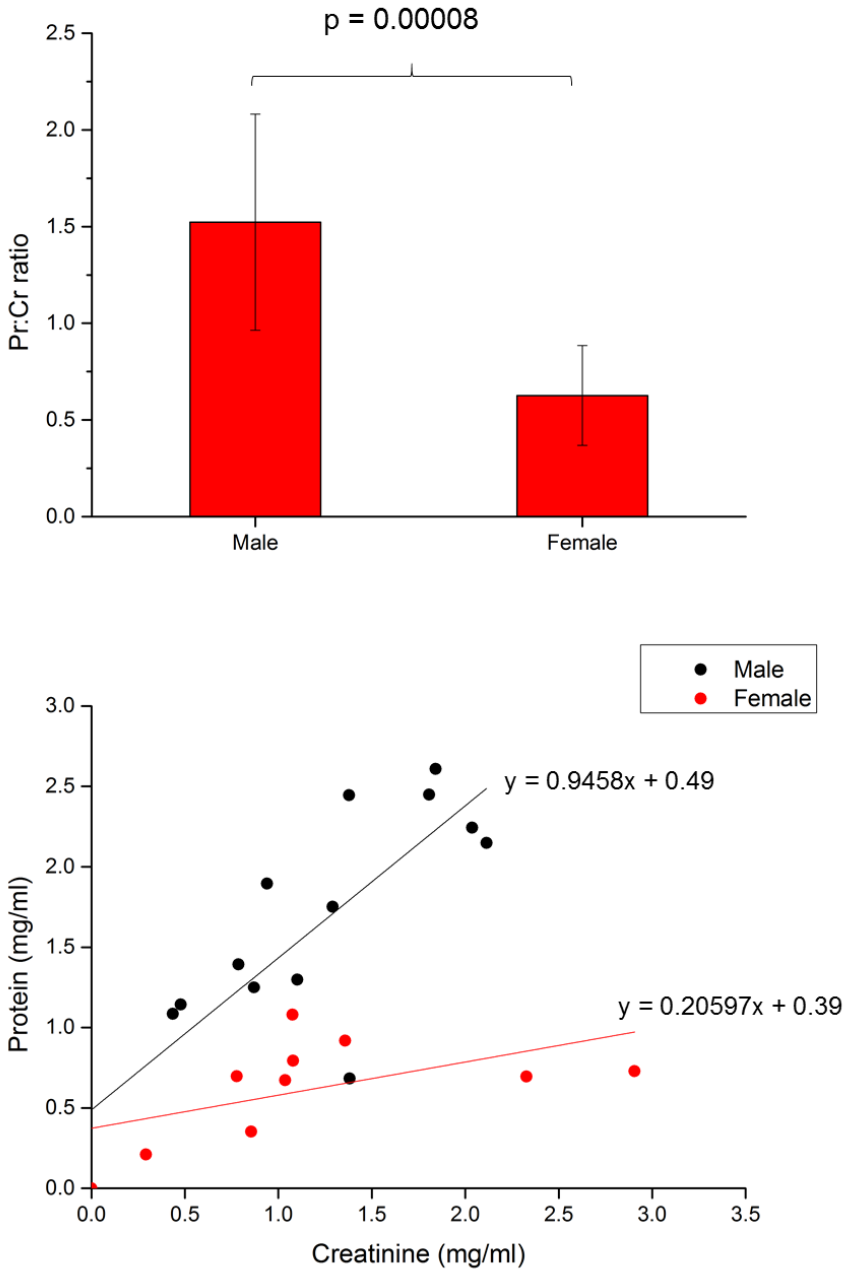


Figure 3.4 Determination of protein concentration, creatinine concentration and protein:creatinine values in male and female *M. spicilegus* urine.
Protein concentration was determined using a Coomassie Plus protein assay kit. A standard curve (0-50 µg/ml) was prepared using BSA. Samples were diluted appropriately in milliQ water, and absorbance readings were measured at 620 nm. Creatinine concentration was measured using a creatinine assay kit. A creatinine standard curve (0-30 µg/ml) was prepared and samples diluted appropriately in milliQ water. Absorbance readings were measured at 570 mm. Top panel: protein: creatinine ratio calculations to correct for urine dilution (error bars ± SD: male (n=13) female (n=9)). The difference between the ratios in male and female urine was statistically significant (Welch t-test, t = 5.06; df = 17.97; p = 8.2×10^{-5}). Bottom panel: calculated protein and creatinine concentrations.

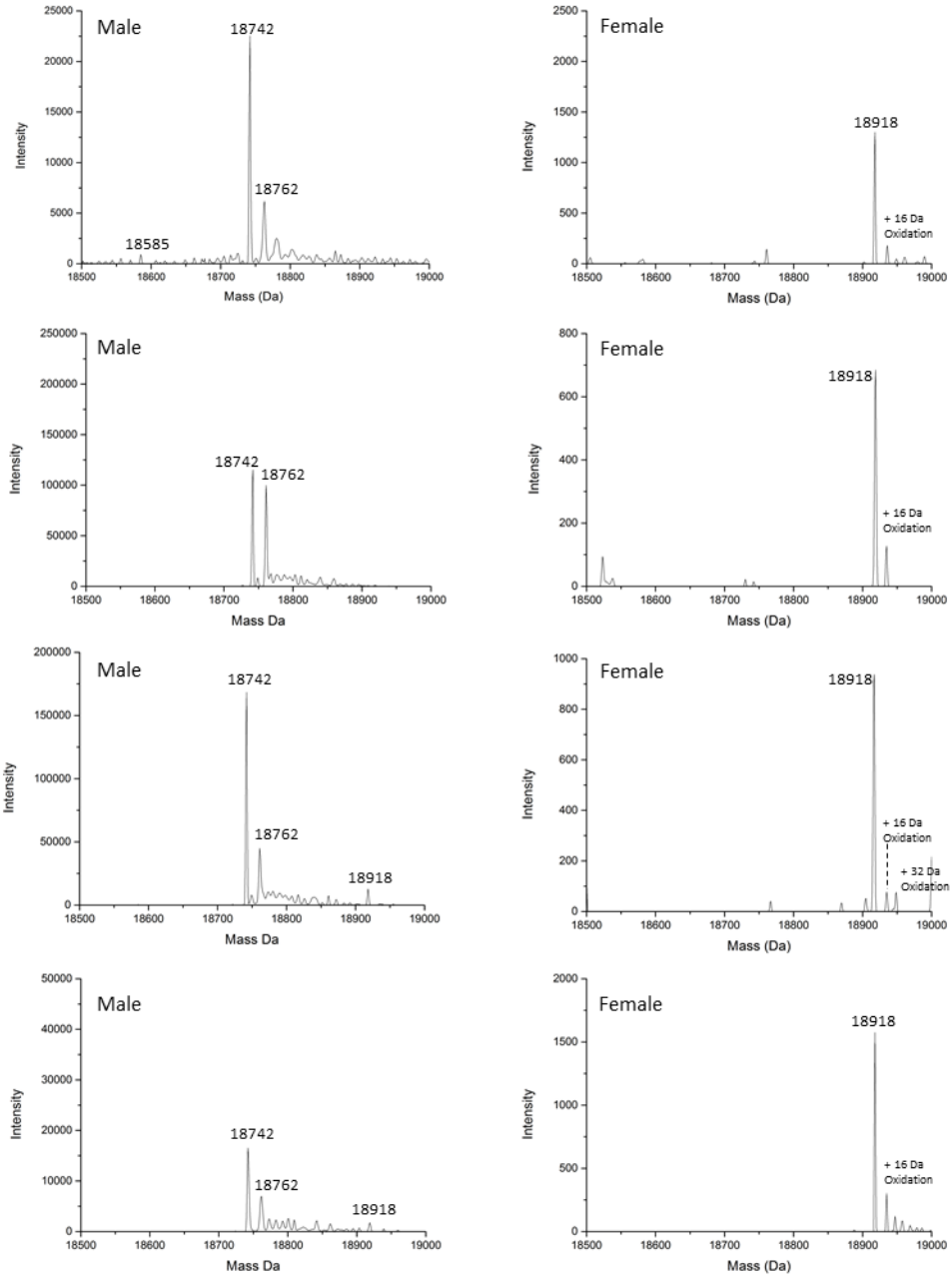


Figure 3.5 Determination of an accurate molecular weight of the ~ 18 kDa proteins in male and female *M. spicilegus* by ESI-MS.
Urine samples from male and female *M. spicilegus* were desalted as described in Chapter 2 and diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). In female urine samples, there is a peak that is + 16 Da of the 18918 Da peak. This mass increase is likely to be due to the oxidation of a methionine residue in the 18918 Da protein.

with known *M. m domesticus* MUPs highlighted many similarities between the proteins in male urine and central *M. m. domesticus* MUPs, but several endoprotease Lys-C peptide masses present in all male *M. spicilegus* samples did not match any known MUP peptide masses (Figures 3.6 and 3.7). In the two major bands selected from the male SDS-PAGE analysis, many peptides of the same mass were observed in both bands. For example, two masses identified by database searching as central MUP peptides, 1567 Da and 1596 Da, were seen in the spectra for both male bands, along with the unidentified masses 1139 Da and 1857 Da (Figures 3.6, 3.7 and 3.9). There were differences in the PMFs obtained for the two male bands, for example the central MUP peptide 2592 Da was present in just the PMF for the lower, less intense male protein band, and the unidentified peptide mass of 1450 Da was only present in the more intense, upper male band (Figures 3.6 and 3.7). Peptide mass fingerprints from female samples had fewer peptides, with most of them matching *M. m. domesticus* peripheral MUP masses (Figures 3.8 and 3.9). The peak at 1613 Da (+ 16 Da of the 1597 Da MUP 6 peptide), which could result from the oxidation of a methionine residue in the peptide sequence, confers with the + 16 Da mass shift in the intact mass analysis of female urine (Figure 3.5). There were no peptide mass similarities between the female and male bands apart from the unidentified mass of 1450 Da, which was seen in the female and the larger, upper male band (Figures 3.6, 3.8 and 3.9). In the female band, this was the only high-intensity observed mass that was not matched to peripheral MUP peptide masses via database searching.

The fact that numerous peptide peaks in the PMF spectra matched masses of peptides generated from Lys-C digestion of known house mouse MUPs provided presumptive evidence for these *M. spicilegus* urinary proteins also being MUPs.

3.2.4 Protein discovery by LC-MS

Male and female urine samples were digested in-solution with endoprotease Lys-C as described in Chapter 2, and the resulting digested material was analysed using an Ultimate 3000 nano system coupled to a QExactive mass spectrometer. PEAKS 7 software was used to process the raw data. Peptides (and the proteins from which they were generated) were identified by searching against a Uniprot mouse protein database. The database search parameters used were those described in Chapter 2.

In the male urine sample, PEAKS software identified 740 peptide-spectrum matches

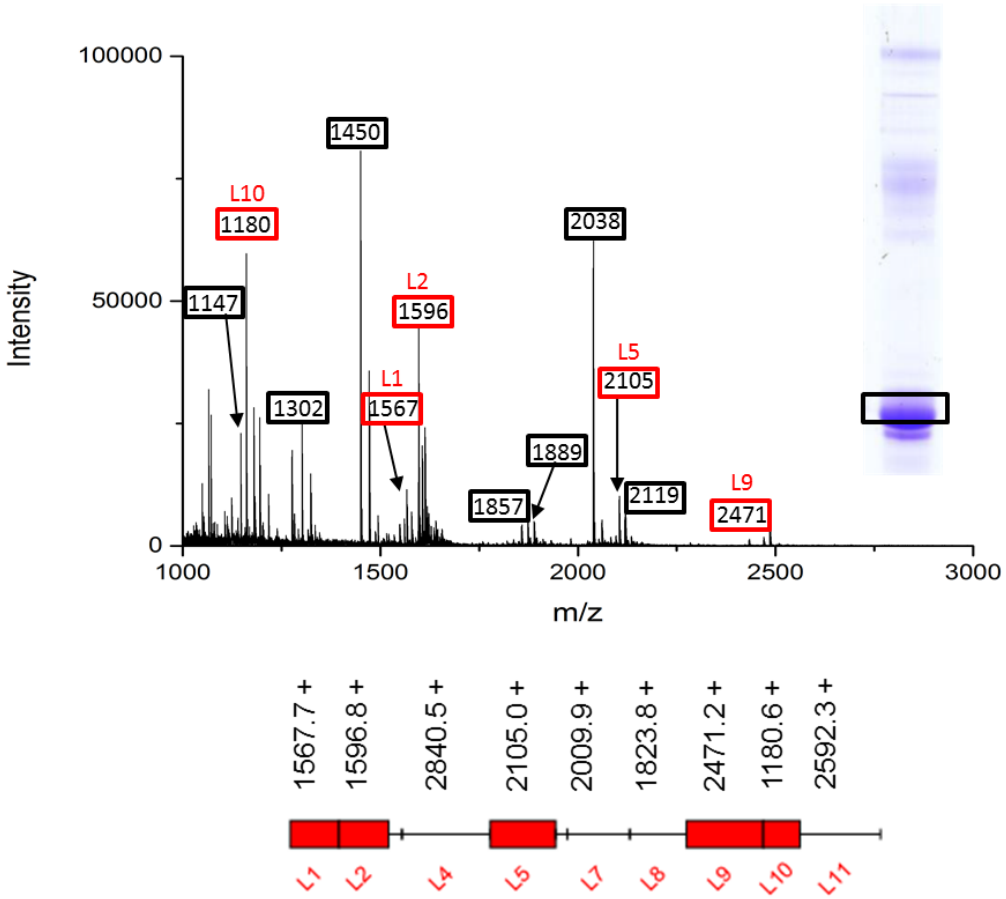


Figure 3.6 Peptide mass fingerprinting of male *M. spicilegus* urine sample.
Pieces of gel were extracted from the protein bands of interest on the SDS-PAGE gels (Figure 3.3), subjected to in-gel digestion and resulting peptides were analysed using MALDI-TOF-MS. The spectra were searched against the MUPS_mat database using Mascot. The digested protein band from male urine showed peptides that matched MUP 7 peptides (highlighted in red), the peptide mass fingerprint of which is shown below the spectrum. Novel peptides are highlighted in black boxes.

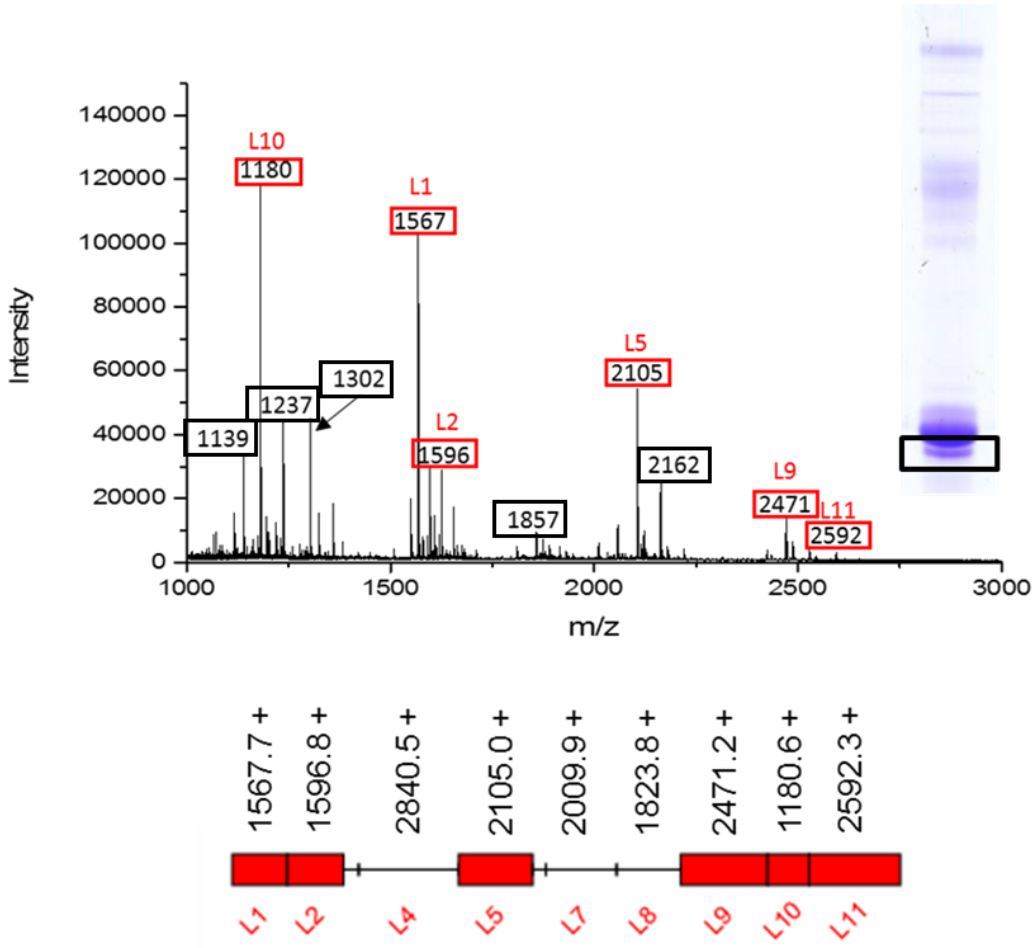


Figure 3.7 Peptide mass fingerprinting of male *M. spicilegus* urine sample.
Pieces of gel were extracted from the protein bands of interest on the SDS-PAGE gels (Figure 3.3), subjected to in-gel digestion and resulting peptides were analysed using MALDI-TOF-MS. The spectra were searched against the MUPS_mat database using Mascot. The digested protein band from male urine showed peptides that matched MUP 7 peptides (highlighted in red), the peptide mass fingerprint of which is shown below the spectrum. Novel peptides are highlighted in black boxes.

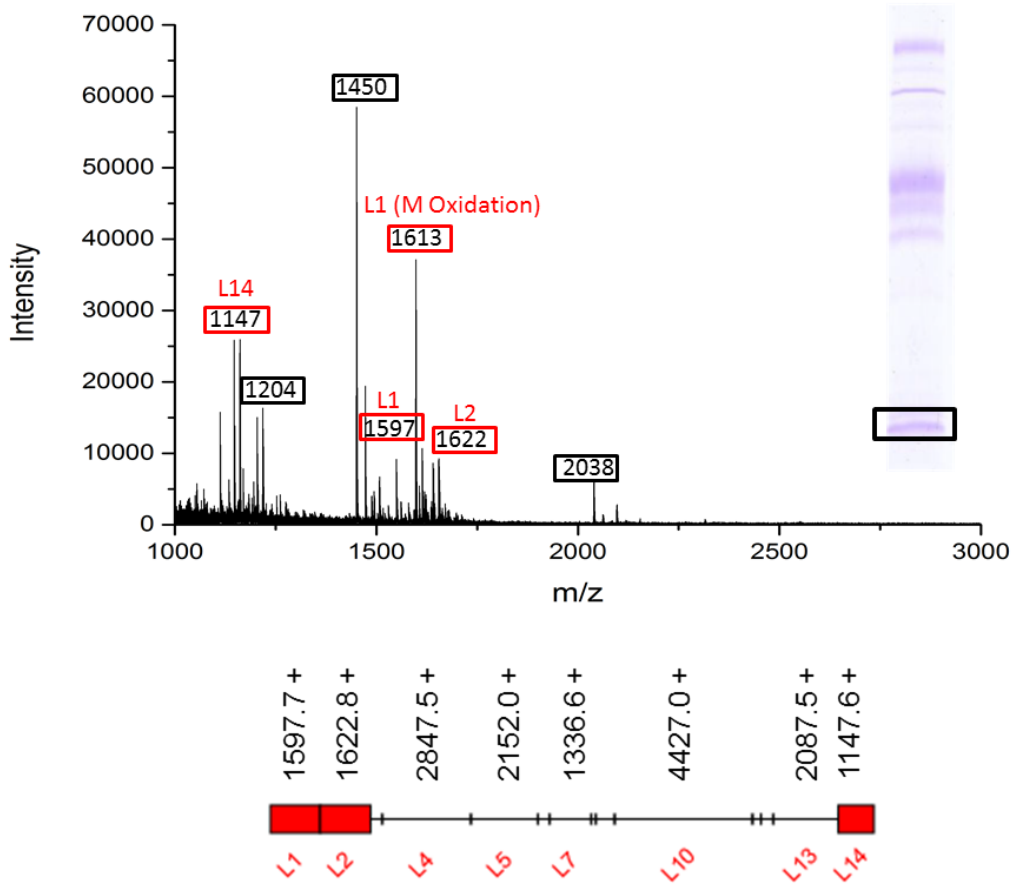


Figure 3.8 Peptide mass fingerprinting of female *M. spicilegus* urine sample.
Pieces of gel were extracted from the protein band of interest on the SDS-PAGE gels (Figure 3.3), subjected to in-gel digestion and resulting peptides were analysed using MALDI-TOF-MS. The spectra were searched against the MUPS_mat database using Mascot. The digested protein band from female urine showed peptides that matched MUP 6 peptides (highlighted in red), the peptide mass fingerprint of which is shown below the spectrum. Novel peptides are highlighted in black boxes.

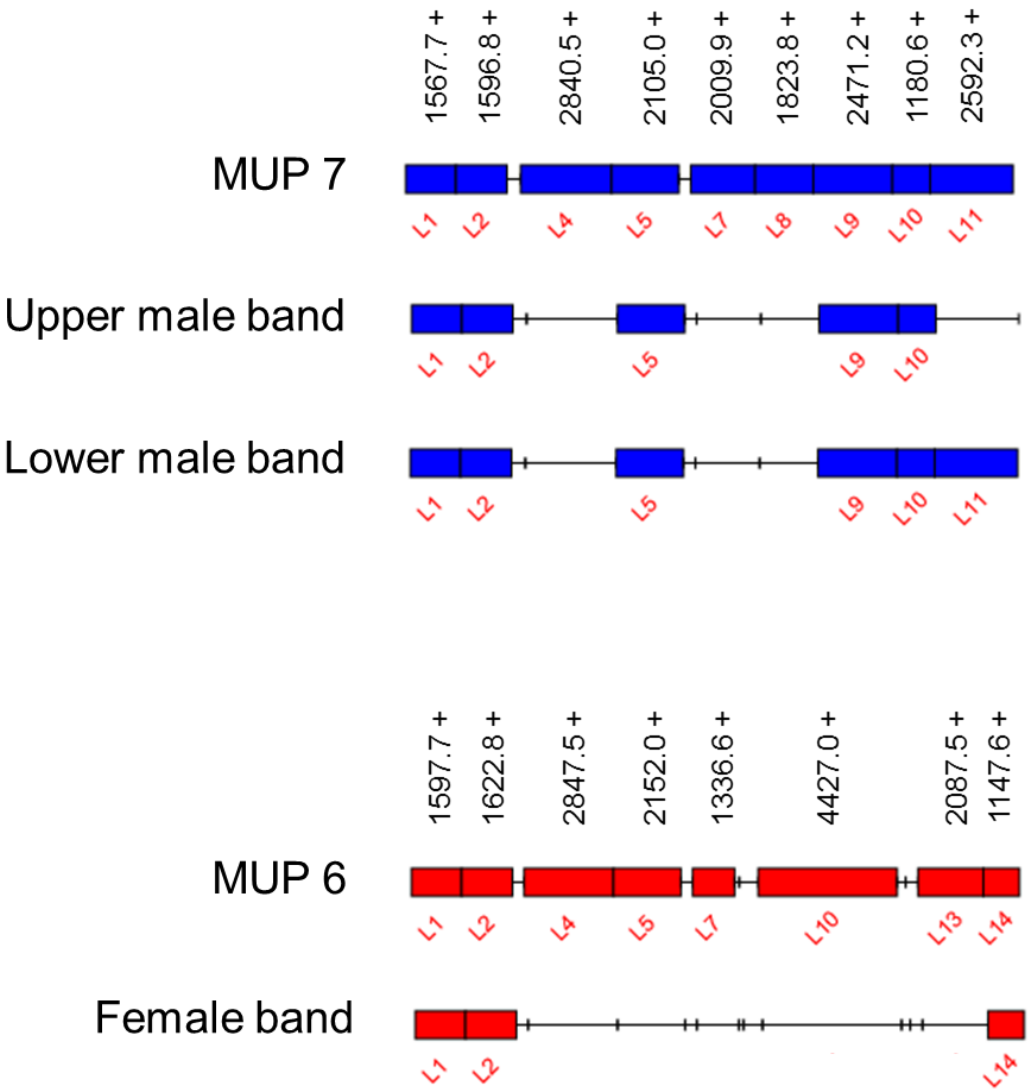


Figure 3.9 Identified MUP peptides in male and female *M. spicilegus* urine samples. Pieces of gel were extracted from the protein band of interest on the SDS-PAGE gels (Figure 3.3), subjected to in-gel digestion and resulting peptides were analysed using MALDI-ToF-MS. The spectra (Figures 3.6 - 3.8) were searched against the MUPS_mat database using Mascot. The digested protein bands from male urine resulted in peptides that matched MUP 7 peptides (indicated in blue). The digested protein band from female urine resulted in peptides that matched MUP 6 peptides (indicated in red).

and 116 proteins. The top 20 identified proteins, along with the number of peptides identified, protein coverage and the presence of any post-translational modifications (PTMs), are outlined in Table 3.1. Of great interest is that the top two identified proteins are both MUPs – the top identified protein is central MUP 1 – with 72% protein sequence coverage and 25 identified peptides, closely followed by peripheral MUP 6, with the same sequence coverage and 24 identified peptides. Also identified are MUPs 3, 4, 5 and 20, all of which are peripheral MUPs.

In female urine, PEAKS software identified 441 peptide-spectrum matches and 83 proteins. The top 20 identified proteins are outlined in Table 3.2. In general, the proteins identified in female urine are similar to those identified in male urine – however, only peripheral MUPs 3 and 4 were in the top 20 identified proteins, with MUPs 5 and 20 (peripheral) also being identified, with only 18% and 20% of the protein sequence covered, respectively.

Thus, predicted from SDS-PAGE, ESI-MS and MALDI-ToF-MS analysis, and confirmed by LC-MS/MS, *M. spicilegus* urine does contain MUPs that are similar to those seen in *M. m. domesticus* urine. However, the fact that the molecular weights of the identified *M. spicilegus* MUPs do not match to the molecular weights of any known *M. m. domesticus* MUPs, and that LC-MS/MS analysis indicates less than full sequence coverage of any known MUPs, means that *M. spicilegus* MUPs are different in sequence to known *M. m. domesticus* MUPs. To identify the sequences of each *M. spicilegus* MUP, each protein was separated from other urinary constituents and digested, and the peptides resulting from proteolysis were sequenced *de novo* using MS/MS spectra generated from LC-MS/MS analyses.

3.2.5 Separation of MUPs using strong anion exchange (SAX) chromatography

To associate identified sequences to specific male *M. spicilegus* MUPs, strong anion exchange (SAX) chromatography was used to separate the proteins in male urine. The buffers used and the pH conditions of the analysis were chosen based on the isoelectric point (pI) of previously studied MUPs – as MUPs have a pI of around 3.2 – 3.9 (Cheetham *et al.* 2009), a buffer pH of 8.8 was selected to confer the proteins with the negative charge necessary for anion exchange chromatography (Robertson *et al*. 1996).

Whole urine was desalted and separated into fractions using SAX chromatography with UV detection at 280 nm. Fraction collection was manual to ensure that each peak

Table 3.1 A list of the proteins identified in male *M. spicilegus* urine from PEAKS database search.
Raw data was processed using PEAKS software and spectra were searched against a Uniprot mouse protein database. PEAKS identified a large number of peptides matching to MUP peptides. MUPs are highlighted in red.

Accession	Protein description	Protein coverage (%)	No. of peptides	PTM(s)
P11588	Major urinary protein 1	72	25	Carbamidomethylation; Oxidation (M)
P02762	Major urinary protein 6	72	24	Carbamidomethylation; Oxidation (M)
P07724	Serum albumin	50	26	Carbamidomethylation
P01132	Pro-epidermal growth factor	20	21	Carbamidomethylation; Oxidation (M)
Q91X17	Uromodulin	17	12	Carbamidomethylation; Oxidation (M)
P11590	Major urinary protein 4	62	12	Carbamidomethylation
P04939	Major urinary protein 3	36	11	Carbamidomethylation
Q61646	Haptoglobin	23	7	Carbamidomethylation
Q61838	Alpha-2-macroglobulin	8	12	
Q00897	Alpha-1-antitrypsin 1-4	26	10	
P11087	Collagen alpha-1(I) chain	5	9	Carbamidomethylation
P11591	Major urinary protein 5	36	8	Carbamidomethylation
Q5FW60	Major urinary protein 20	30	9	
O09043	Napsin-A	12	4	Carbamidomethylation
P51910	Apolipoprotein D	39	6	Carbamidomethylation; Oxidation (M)
A2ARV4	Low-density lipoprotein receptor-related protein 2	2	10	Carbamidomethylation
P15947	Kallikrein-1	12	4	Carbamidomethylation
Q60648	Ganglioside GM2 activator	32	4	Carbamidomethylation
Q9DAU7	WAP four-disulfide core domain protein 2	19	3	Carbamidomethylation
P07758	Alpha-1-antitrypsin 1-1	18	8	

Table 3.2 A list of the proteins identified in female *M. spicilegus* urine from PEAKS database search.
Raw data was processed using PEAKS software and spectra were searched against a Uniprot mouse protein database. PEAKS identified a number of peptides matching to MUP peptides. MUPs are highlighted in red.

Accession	Protein description	Protein coverage (%)	No. of peptides	PTM(s)
P07724	Serum albumin	48	24	Carbamidomethylation; Oxidation (M)
Q60574	Haptoglobin	37	12	Carbamidomethylation; Oxidation (M)
P07759	Serine protease inhibitor A3K	25	9	Oxidation (M)
Q61646	Haptoglobin	34	11	Carbamidomethylation; Oxidation (M)
P15947	Kallikrein-1	41	8	Carbamidomethylation; Oxidation (M)
P01132	Pro-epidermal growth factor	14	11	Carbamidomethylation; Oxidation (M)
P07758	Alpha-1-antitrypsin 1-1	25	9	Oxidation (M)
Q00896	Alpha-1-antitrypsin 1-3	25	9	Oxidation (M)
Q921I1	Serotransferrin	22	12	Carbamidomethylation
Q61838	Alpha-2-macroglobulin	12	14	Carbamidomethylation
O70570	Polymeric immunoglobulin receptor	11	6	Carbamidomethylation
Q07456	Protein AMBP	23	5	Carbamidomethylation
P22599	Alpha-1-antitrypsin 1-2	21	8	
Q00897	Alpha-1-antitrypsin 1-4	21	8	
P51910	Apolipoprotein D	36	7	Carbamidomethylation
P11590	Major urinary protein 4	42	6	Carbamidomethylation
Q9DAU7	WAP four-disulfide core domain protein 2	37	6	Carbamidomethylation; Oxidation (M)
P04939	Major urinary protein 3	33	4	Carbamidomethylation; Oxidation (M)
O09043	Napsin-A	13	2	Carbamidomethylation
P23953	Liver carboxylesterase N	9	4	Carbamidomethylation

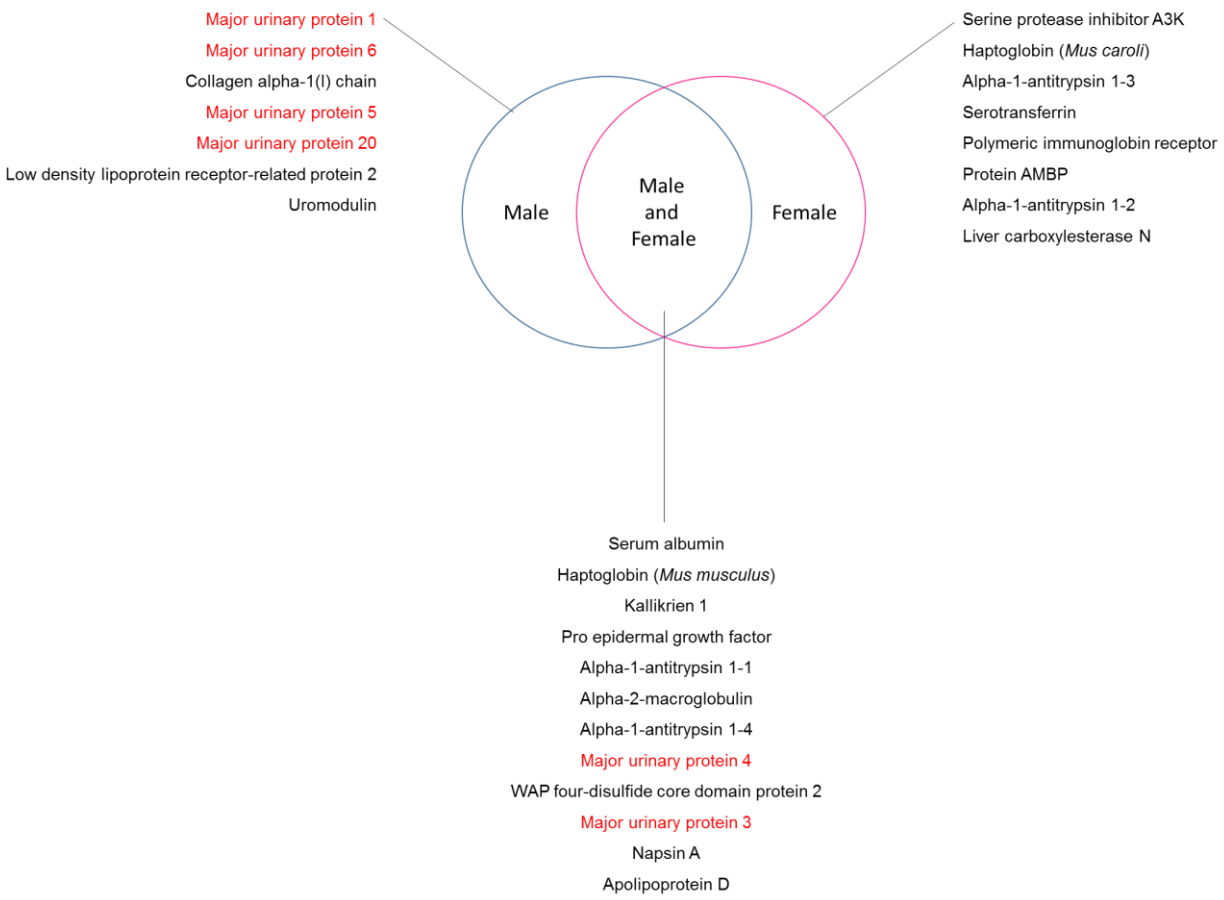


Figure 3.10 The proteins identified in *M. spicilegus* urine.
Raw data was processed using PEAKS software and spectra were searched against a Uniprot mouse protein database. PEAKS identified proteins that were present in male urine only, proteins that were present in female urine only, and proteins that were present in both. Major urinary proteins are highlighted in red.

at 280 nm was collected in a single fraction. Each fraction was desalted and diluted 1 in 10 in 0.1% (v/v) formic acid before being subjected to ESI-MS analysis to determine the MUP(s) present in each fraction. The 18585 Da MUP could not however be resolved from the other two male MUPs (Figure 3.11). Fraction 1 contained the 18742 Da MUP only, Fraction 3 contained the 18762 Da MUP only, and Fraction 2 contained proteins of masses 18742 Da, 18762 Da and 18585 Da. Due to the lower complexity in protein expression and lower abundance of protein in female urine (confirmed by SDS-PAGE and intact mass analysis, Figures 3.3 and 3.5), fractionation prior to protein sequencing analysis was not performed.

3.2.6 Sequencing of *Mus spicilegus* MUPs *de novo*

In order to determine full protein sequences for each male *M. spicilegus* MUP observed, the MUP fractions collected from the SAX chromatographic separation were desalted and separately digested with endoproteases Lys-C and Glu-C. The 18742 Da and 18762 Da MUPs that were completely isolated in their SAX fractions were sequenced first, so the peptides present in Fraction 2 that were not part of either of these sequences would be peptides from 18585 Da MUP. This, supported by the intact protein mass measurement, would enable the definitive sequencing of each male *M. spicilegus* MUP. To sequence the 18918 Da MUP, desalted and concentrated female urine was separately digested with each endoprotease. After proteolysis, all samples were subjected to LC-MS/MS analysis using the Thermo QExactive.

Raw data were processed using PEAKS software, which sequences the identified peptides *de novo*. MS/MS analysis yielded complete and accurate sequence information for the majority of the *M. spicilegus* MUP peptides; in all but one case, overlapping sequence information solved sequence ambiguities. Most sequences were above the 85% confidence threshold (Supplementary Material), and MS/MS spectra were good quality. All peptide sequences were searched against a database of *Mus* proteins using BLASTp, an online local alignment search tool, to determine any similarities or differences between the *M. spicilegus* MUP peptides and known mouse MUPS. BLAST search results determined that the peptides from proteins specifically in male *M. spicilegus* urine (18585 Da, 18742 Da and 18762 Da) were highly homologous in sequence to *M. m. domesticus* central MUPs, and the non-sex specific protein 18918 Da in *M. spicilegus* was homologous to *M. m. domesticus* peripheral MUP 6 (Tables 3.3 and 3.4).

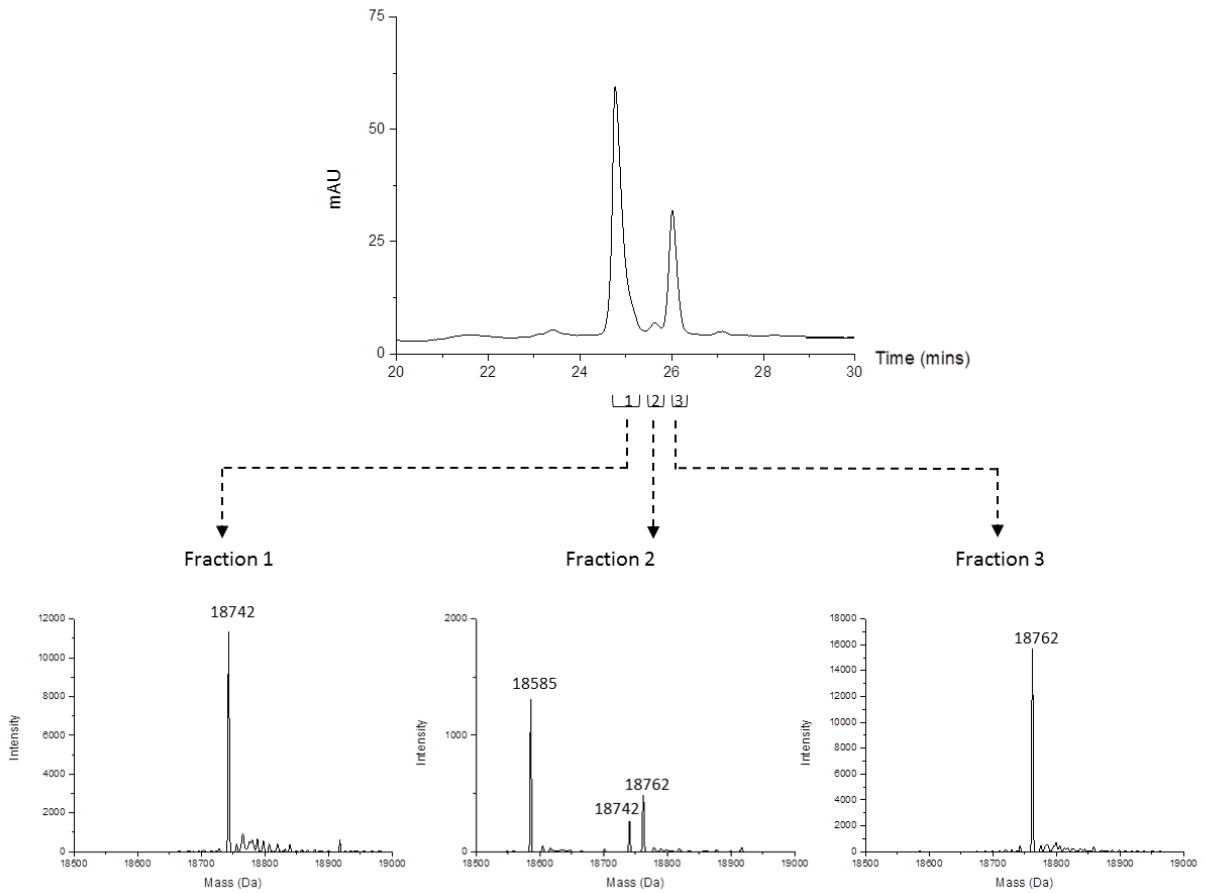


Figure 3.11 ESI-MS analysis of anion exchange separated proteins from male *M. spicilegus* urine.
Desalted urinary proteins from one male were separated using strong anion exchange chromatography. Protein elution was monitored by UV absorbance, and fractions were manually collected during protein elution. The masses of the proteins present in the anion exchange fractions were determined by ESI-MS. MUP mass error was +/- 2 Da.

Table 3.3 BLAST results from male-specific MUP Lys-C *de novo* sequenced peptides. Peptide sequences were assessed using the BLASTp algorithm. Highlighted in red are the amino acid differences between male-specific *M. spicilegus* MUP and *M. musculus* central MUP peptides. The score of the alignment is assigned by the program, and the E value indicates the significance of the sequence match. Identified peptides were searched for in the previous MALDI-ToF-MS analysis of protein bands from SDS-PAGE of male urine (Figures 3.6 - 3.7).

Peptide mass (Da)	*M. spicilegus* MUP	*M. spicilgus* peptide sequence Known MUP peptide sequence	Protein identification	Sequence identity (%)	Score	E value
1566.7	18585, 18742, 18762	EEASSTGRNFNVEK EEASSTGRNFNVEK	MUP 1 Protein *(M. musculus)*	100	46.4	5×10^{-8}
1595.8	18585, 18742, 18762	INGEWHTIILASDK INGEWHTIILASDK	MUP 1 Protein *(M. musculus)*	100	49.4	4×10^{-8}
2850.3	18585, 18742	IEDNGNFRLFLEHIHVLENSLVLK IEDNGNFRLFLEQIHVLENSLVLK	MUP 1 Protein *(M. musculus)*	96	77.8	2×10^{-17}
2103.9	18742	VHTVRDEECSELSMVADK VHTVRDEECSELSMVADK	MUP 7 Protein *(M. musculus)*	100	61.7	4×10^{-12}
1138.5	18585, 18742	TERAGEYSVK TEKAGEYSVT	MUP 1 Protein *(M. musculus)*	89	27.8	0.37
1301.6	18585, 18742	YDGFNTFTIPK YDGFNTFTIPK	MUP 1 Protein *(M. musculus)*	100	40.1	4×10^{-7}
1856.8	18585, 18742	TDYDNFMMSHLINEK TDYDNFLMAHLINEK	MUP 1 Protein *(M. musculus)*	87	49.4	3×10^{-8}
2470.7	18585, 18742	DGQTFQLMGLYGREPDLSSDIK DGETFQLMGLYGREPDLSSDIK	MUP 1 Protein *(M. musculus)*	95	71.0	2×10^{-15}
1179.5	18585, 18742	ERFAQLCEK ERFAQLCEK	Group 1 Major Urinary Protein *(M. musculus)*	100	33.3	0.003
2592.8	18585, 18742	HGILRENIIDLSNANRCLQARE HGILRENIIDLSNANRCLQARE	Group 1 Major Urinary Protein *(M. musculus)*	100	74.4	2×10^{-17}
2878.3	18762	IEEHGNFRLFLEQIHVLENSLVLK IEEHGNFRLFLEQIHVLENSLVLK	MUP 15 Protein *(M. musculus)*	100	98.0	3×10^{-20}
2121.9	18762	FHLGRDEECSELSMVADK FHTVRDEECSELSMVADK	MUP *(M. musculus)*	89	58.5	1×10^{-10}
2008.9	18762	AGEYSVTYDGFNTFTIPK AGEYSVTYDGFNTFTIPK	MUP 1 Protein *(M. musculus)*	100	61.3	3×10^{-12}
1870.8	18762	TDYDNFMMTHLINEK TDYDNFLMAHLINEK	MUP 1 Protein *(M. musculus)*	87	49.0	4×10^{-8}
2471.7	18762	DGETFQLMGLYGREPDLSSDIK DGETFQLMGLYGREPDLSSDIK	MUP 1 Protein *(M. musculus)*	100	74.0	2×10^{-18}
1506.7	18762	HGILRENIIDLSK HGILRENIIDLSN	Group 1 Major Urinary Protein *(M. musculus)*	92	41.8	5×10^{-6}
1116.5	18762	ANRCLQARE ANRCLQARE	Group 1 Major Urinary Protein *(M. musculus)*	100	32.5	0.006

Table 3.4 BLAST results from the non-sex specific MUP Lys-C *de novo* sequenced peptides.
Peptide sequences were assessed using the BLASTp algorithm. Highlighted in red are the amino acid differences between non sex specific *M. spicilegus* MUP and *M. musculus* peripheral MUP peptides. The score of the alignment is assigned by the program, and the E value indicates the significance of the sequence match. Identified peptides were searched for in the previous MALDI-ToF-MS analysis of protein bands from SDS-PAGE of female urine (Figure 3.8).

Peptide mass (Da)	*M. spicilegus* MUP	*M. spicilgus* peptide sequence Known MUP peptide sequence	Protein identification	Sequence identity (%)	Score	E value
1597.7	18918	EEASSMGRNFNVEK EEASSMGRNFNVEK	Major urinary protein 6 *(M. musculus)*	100	48.1	1×10^{-7}
1622.8	18918	INGEWYTIILASDK INGEWYTIILASDK	Major urinary protein 6 *(M. musculus); Major urinary protein 20 (M. musculus)*	100	49.8	1×10^{-8}
1181.3	18918	CSEIFLVADK CSEIFLVADK	Major urinary protein 4 (*M. musculus*); Major urinary protein 6 *(M. musculus); Major urinary protein 20 (M. musculus)*	100	35.4	0.001
1450.5	18918	AGEYSVTYDGFNK AGEYSVTYDGFKK	Major urinary protein 6 *(M. musculus)*	92	41.8	1×10^{-5}
606.8	18918	FTVLK FTVLK	Major urinary protein 6 *(M. musculus)*	100	18.9	167
1881.2	18918	TDYDNYIMIHLINKK TDYDNYIMIHLINKK	Major urinary protein 3 (*M. musculus*); Major urinary protein 5 *(M. musculus); Major urinary protein 20 (M. musculus)*	100	55.8	2×10^{-10}
1606.8	18918	SLYGREPDLNSDIK SLYGREPDLNSDIK	Major urinary protein 6 *(M. musculus)*	100	47.7	2×10^{-7}
1663.8	18918	YGREPDLNSDIKEK YGREPDLNSDIKEK	Major urinary protein 6 *(M. musculus)*	100	48.6	8×10^{-8}
2039.3	18918	LCEEHGILRENIIDVTK LCEEHGILRENIIDFTK	Major urinary protein 6 *(M. musculus)*	93	49.0	2×10^{-8}
1147.2	18918	TNRCLQARE TNRCLQARE	Major urinary protein 6 *(M. musculus)*	100	32.9	0.008

Each *M. spicilegus* MUP peptide resulting from the analysis had extremely high sequence homology to known MUP peptides – many of the sequenced peptides were identical to those seen in *M. m. domesticus* MUPs, with all sequenced peptides sharing at least 87% sequence identity with known MUP peptides. Many of the *M. spicilegus* peptides that are identical in sequence to known MUP peptides were initially identified in the peptide mass fingerprinting analysis (Figures 3.6 - 3.9), but LC-MS/MS analysis enabled the confirmation of these sequences as well as identifying the unknown peptide masses. The G-X-W sequence motif (where X represents any amino acid) was present in the peptide sequences of all digested *M. spicilegus* MUP fractions. This sequence motif is highly conserved across all lipocalins, including MUPs, reaffirming that these *M. spicilegus* proteins are indeed MUPs. The high sequence homology of all *M. spicilegus* MUP peptides to *M. m. domesticus* MUP peptides allowed the definitive construction of the peptides of each MUP into the correct protein sequence. Reference to the calculated molecular weights of each MUP confirmed that the correct sequences had been assigned. The MS/MS spectra for the Lys-C digested peptides of the most abundant male specific MUP 18742 Da, from Fraction 1 of the SAX chromatography analysis, are displayed in protein sequence order in Figures 3.12 – 3.21. The MS/MS spectra for the Lys-C digested peptides of MUPs 18762 Da, 18585 Da and 18918 Da, along with the spectra for the Glu-C digested peptides of all four *M. spicilegus* MUPs, are displayed in the Supplementary Material. The sequence coverage obtained from both Lys-C and Glu-C digests for each MUP can be found in Figures 3.22 - 3.25.

For the three male specific MUPs (18742 Da, 18762 Da and 18585 Da), the use of two proteases provided overlapping and therefore full definitive sequence information. For the non sex specific MUP 18918 Da, however, a small part of the protein sequence was not covered (8 amino acid residues). Female urine was then digested using a third protease, Asp-N, to try and deduce the missing sequence, but the use of this protease provided no additional sequence information. Since the 18918 Da MUP peptides show high sequence homology to the peripheral *M. musculus* MUP 6, it was possible that the *M. musculus* MUP sequence NGETFQLM (the sequence usually found in the same part of the MUP where the 18918 Da sequence was missing) might complete this MUP sequence. However, placing this sequence in the otherwise fully sequenced 18918 Da MUP resulted in a molecular weight of just 18882 Da for this protein (36 Da less than the calculated molecular weight). The only amino acid substitution in the sequence NGETFQLM that would give rise to a 36 Da increase in molecular weight would be T → H, resulting in the

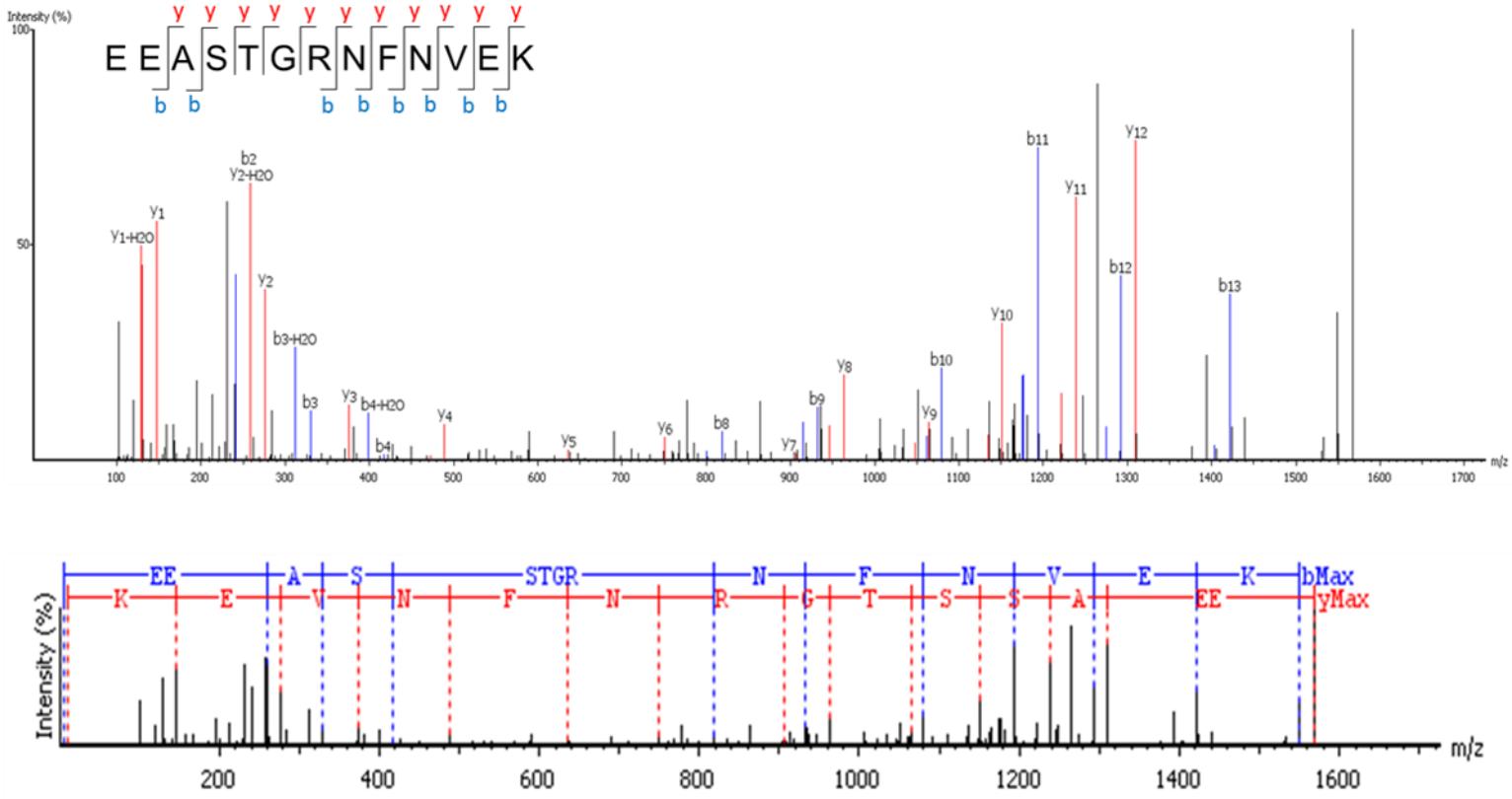


Figure 3.12 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1567 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).

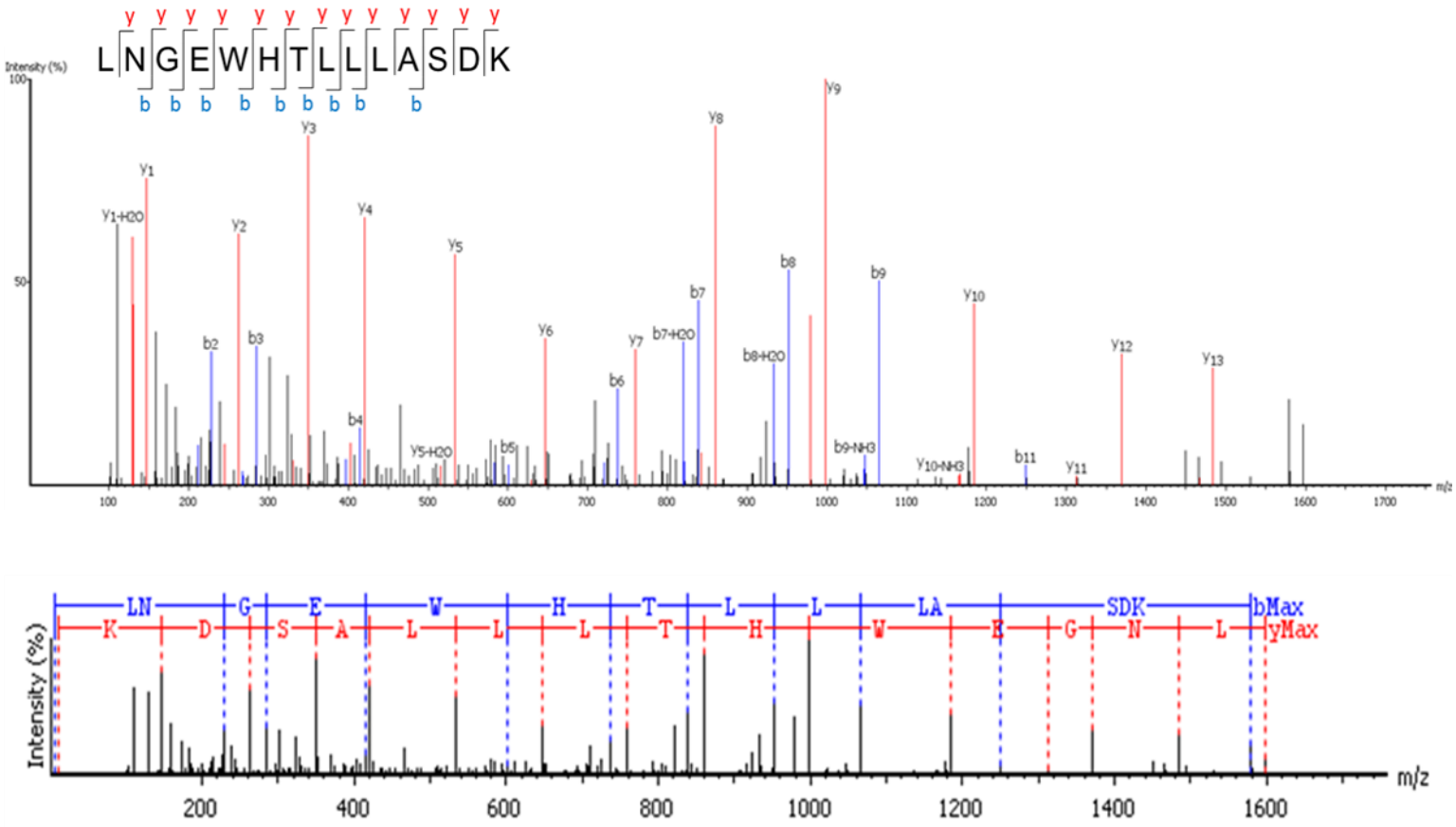


Figure 3.13 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1596 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

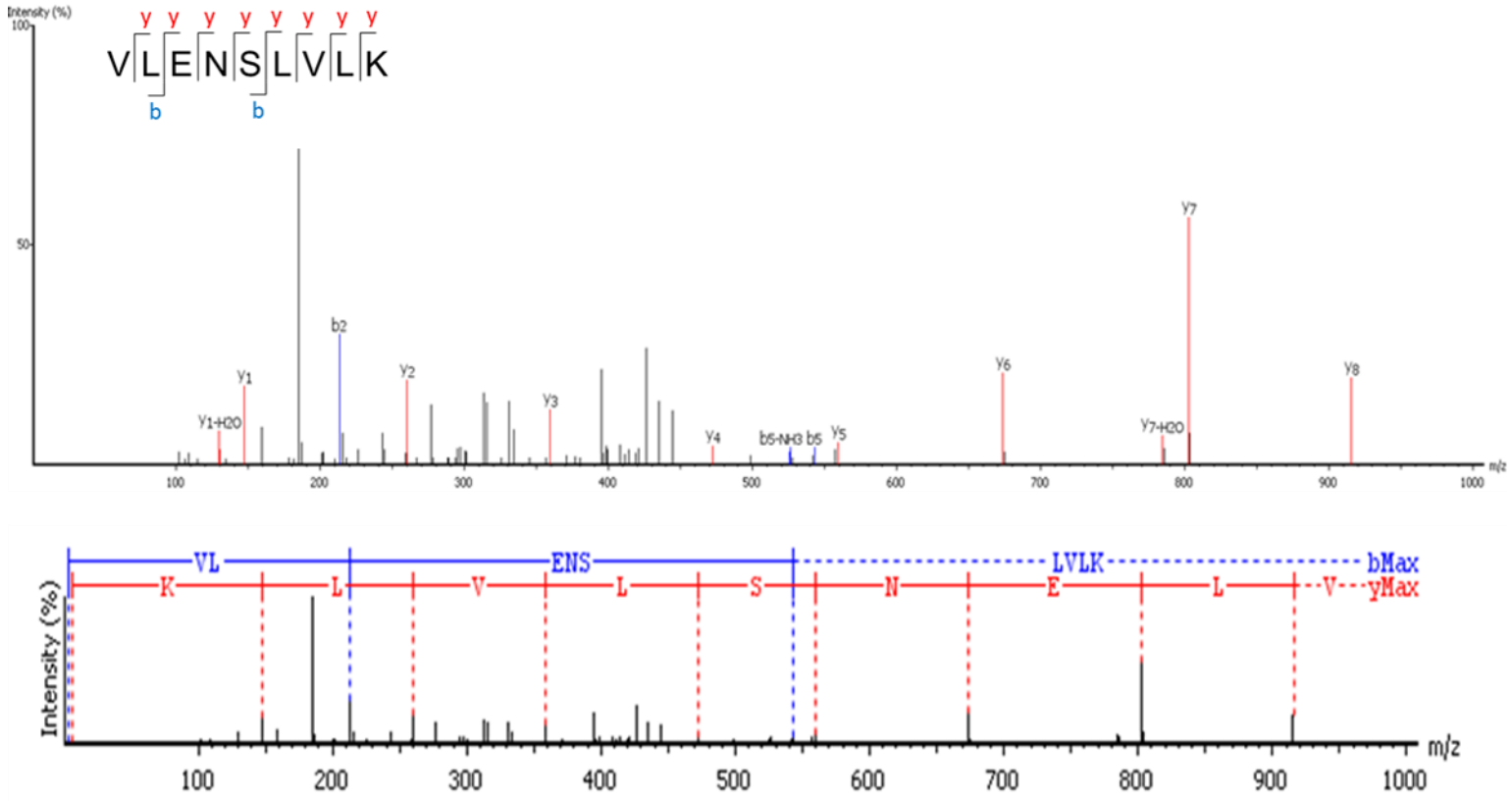


Figure 3.14 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1014 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

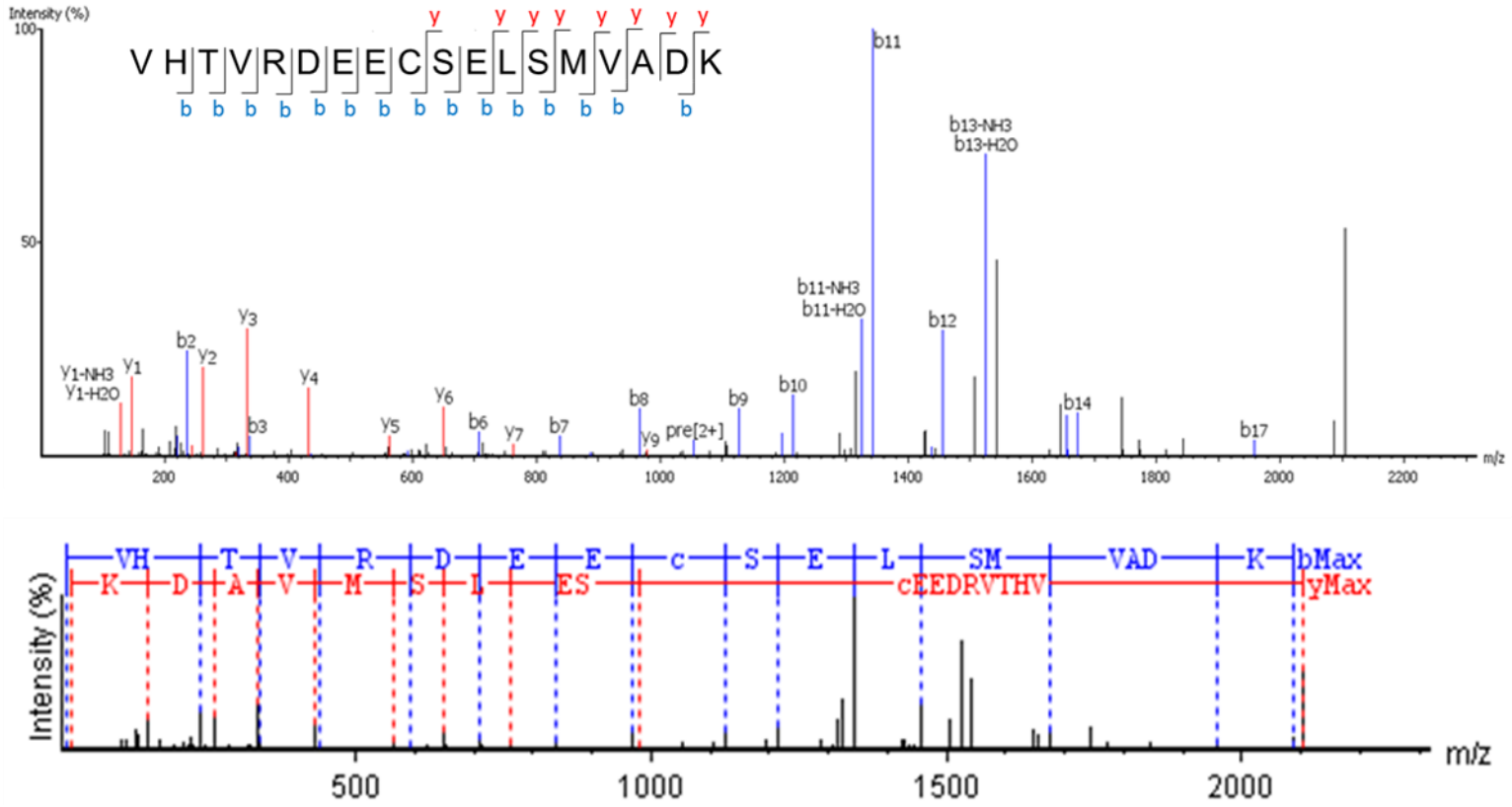


Figure 3.15 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2105 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

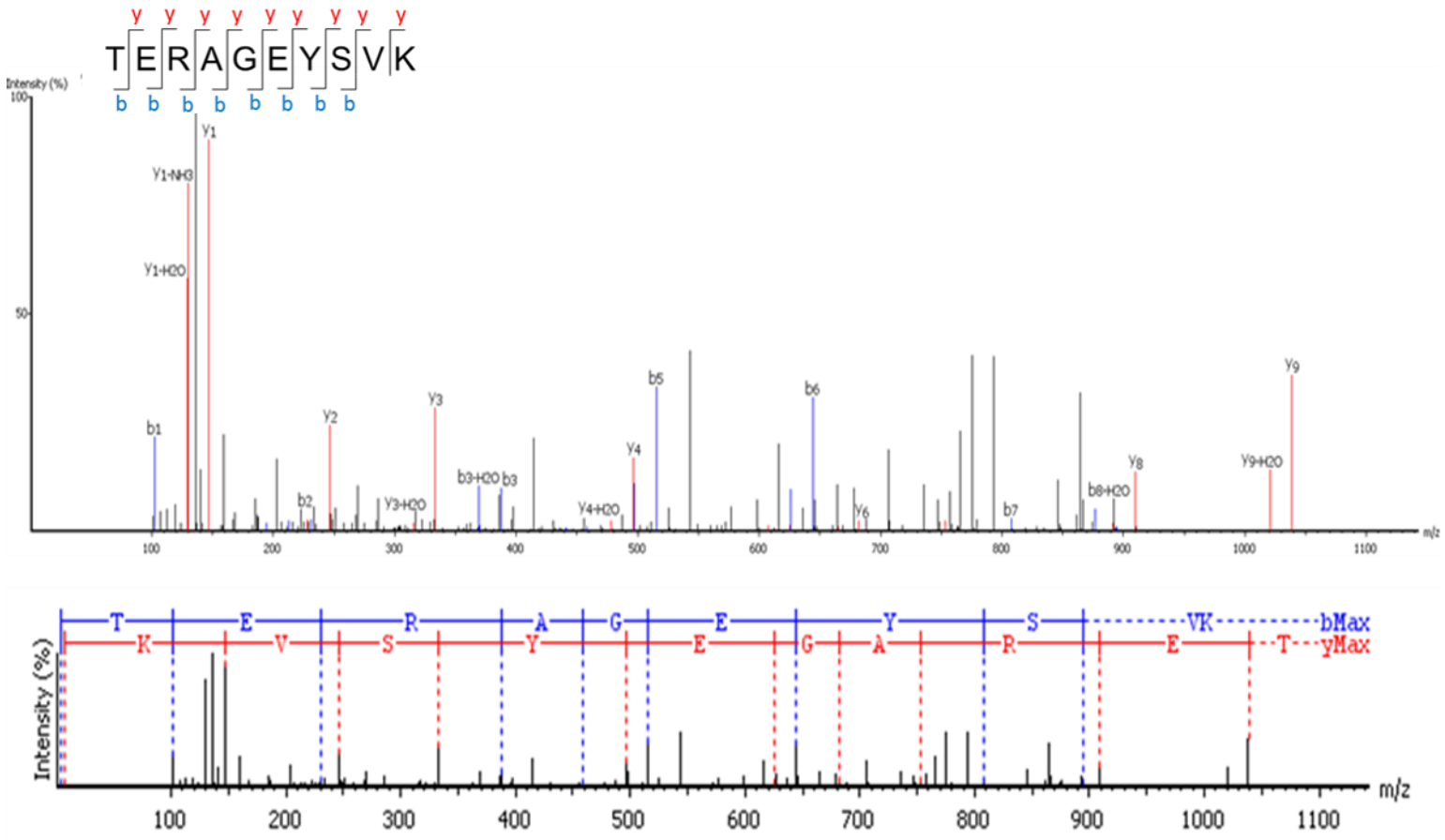


Figure 3.16 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1139 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).

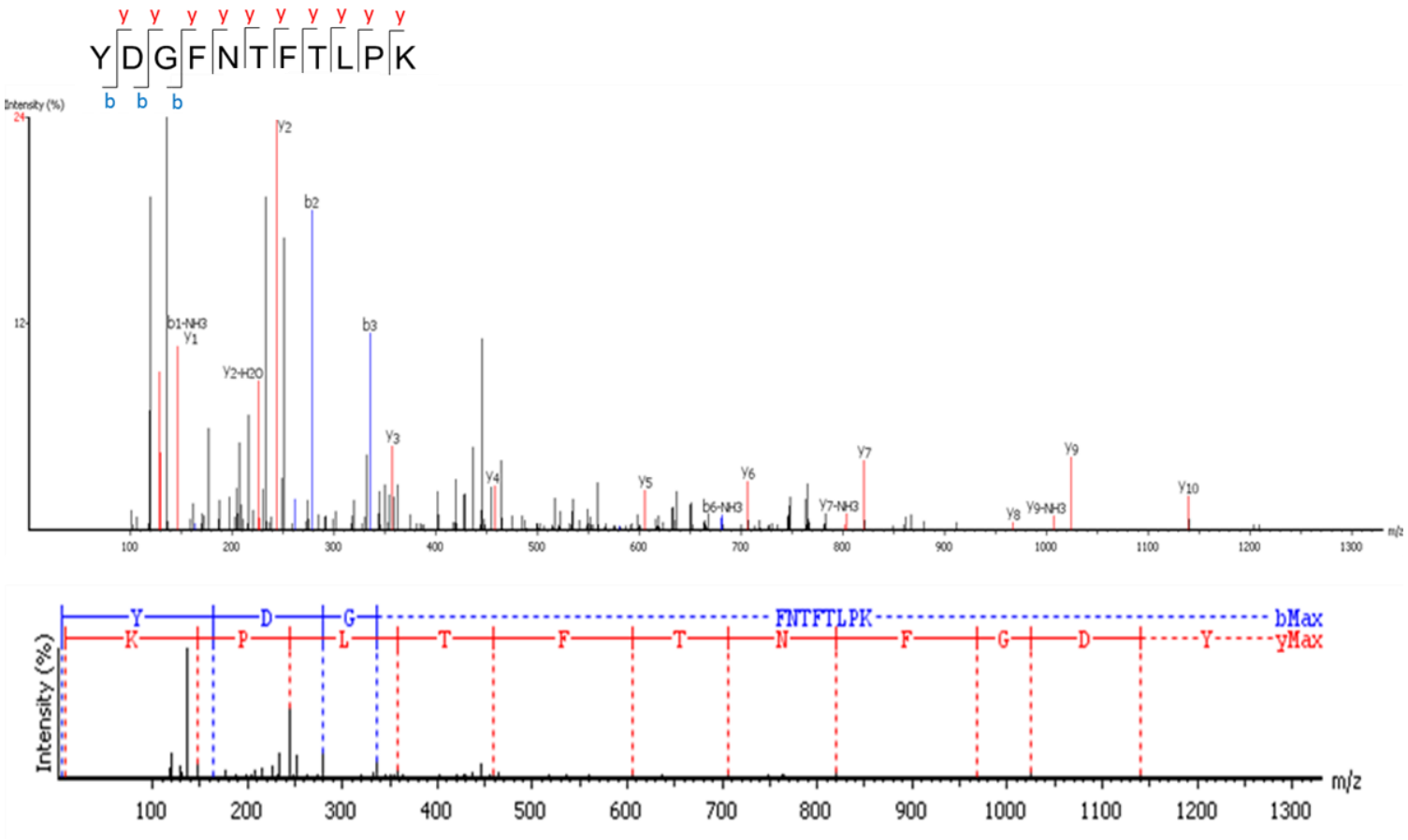


Figure 3.17 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1302 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

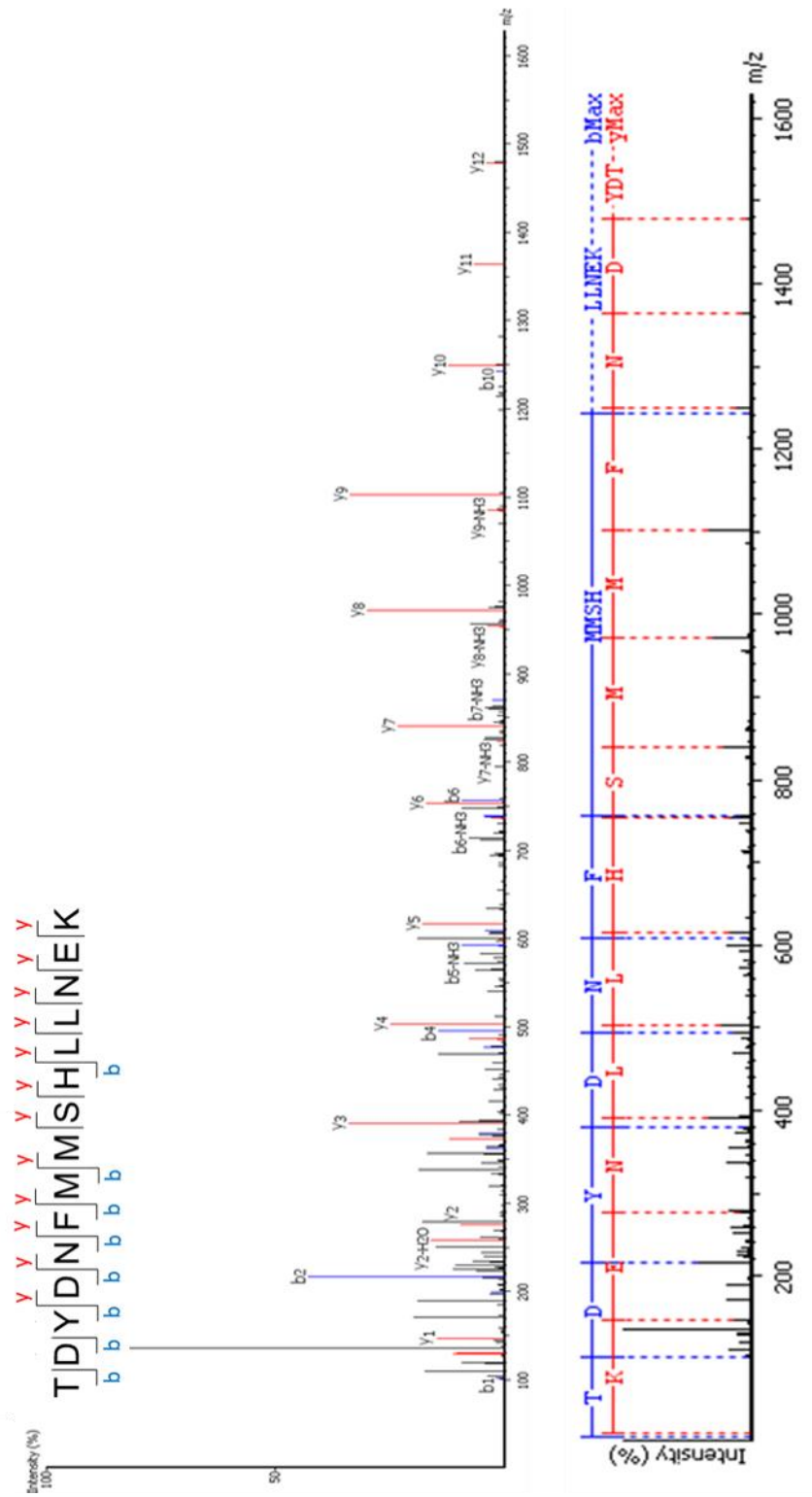


Figure 3.18 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1858 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

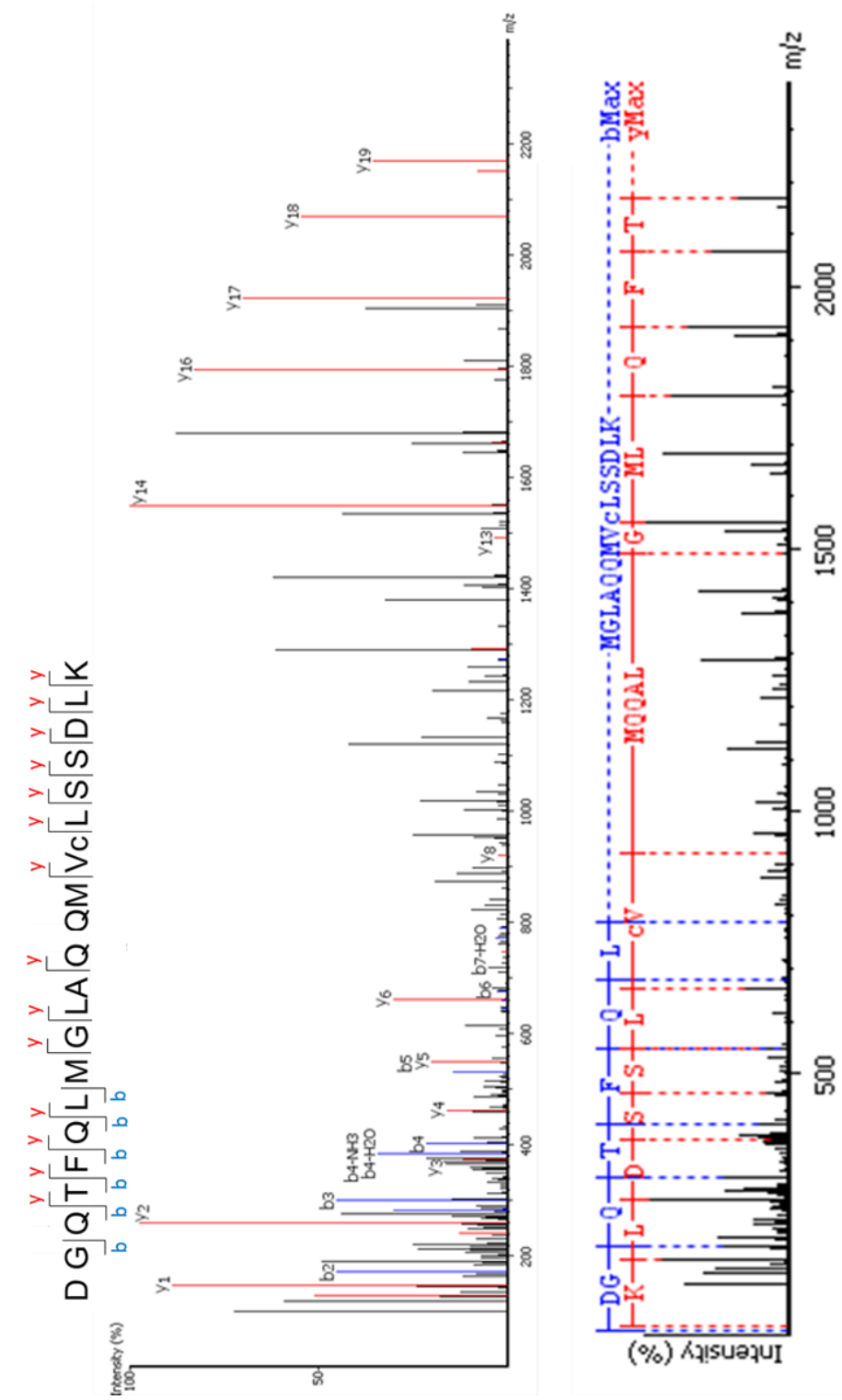


Figure 3.19 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2471 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

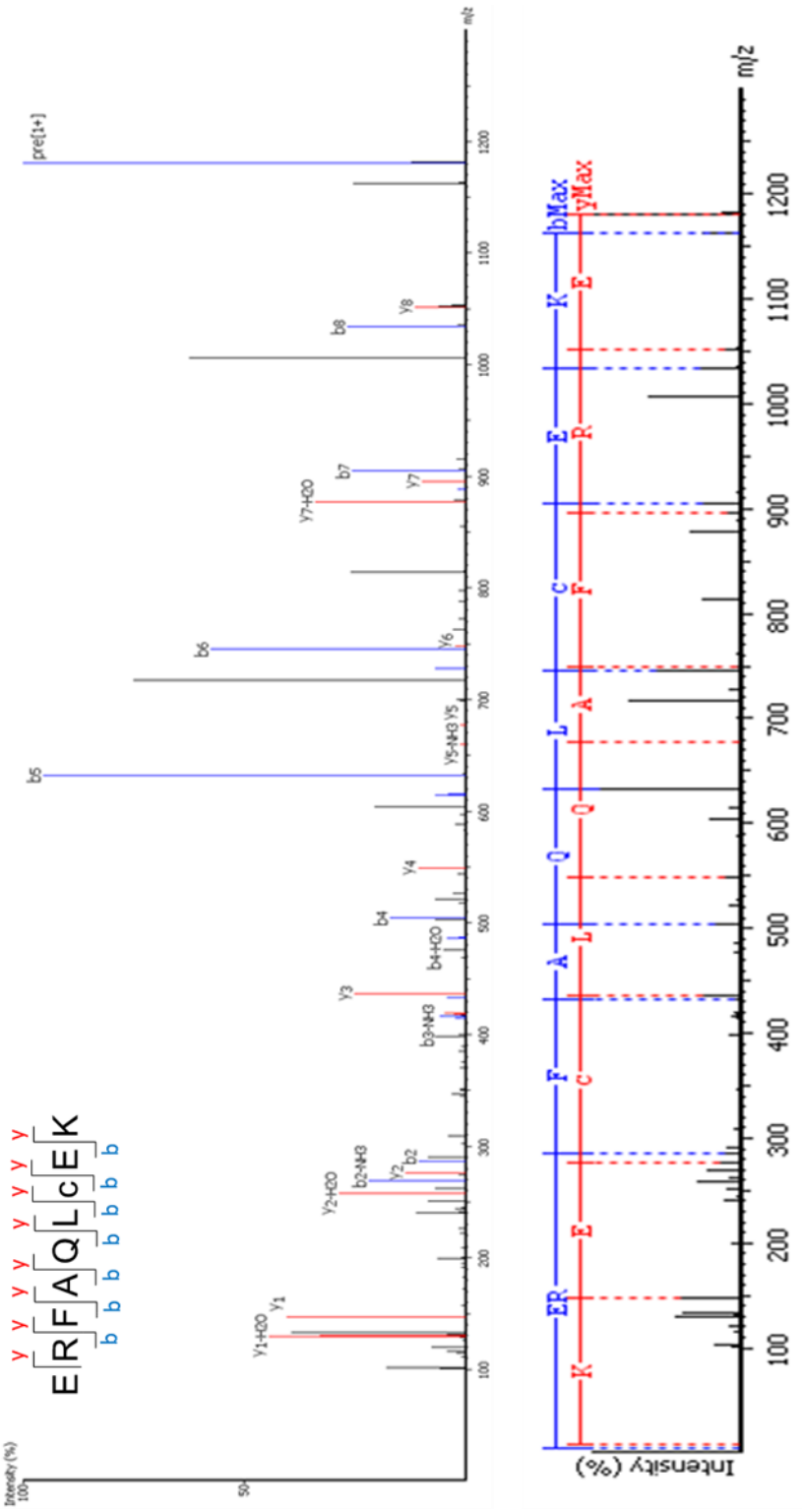


Figure 3.20 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1180 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

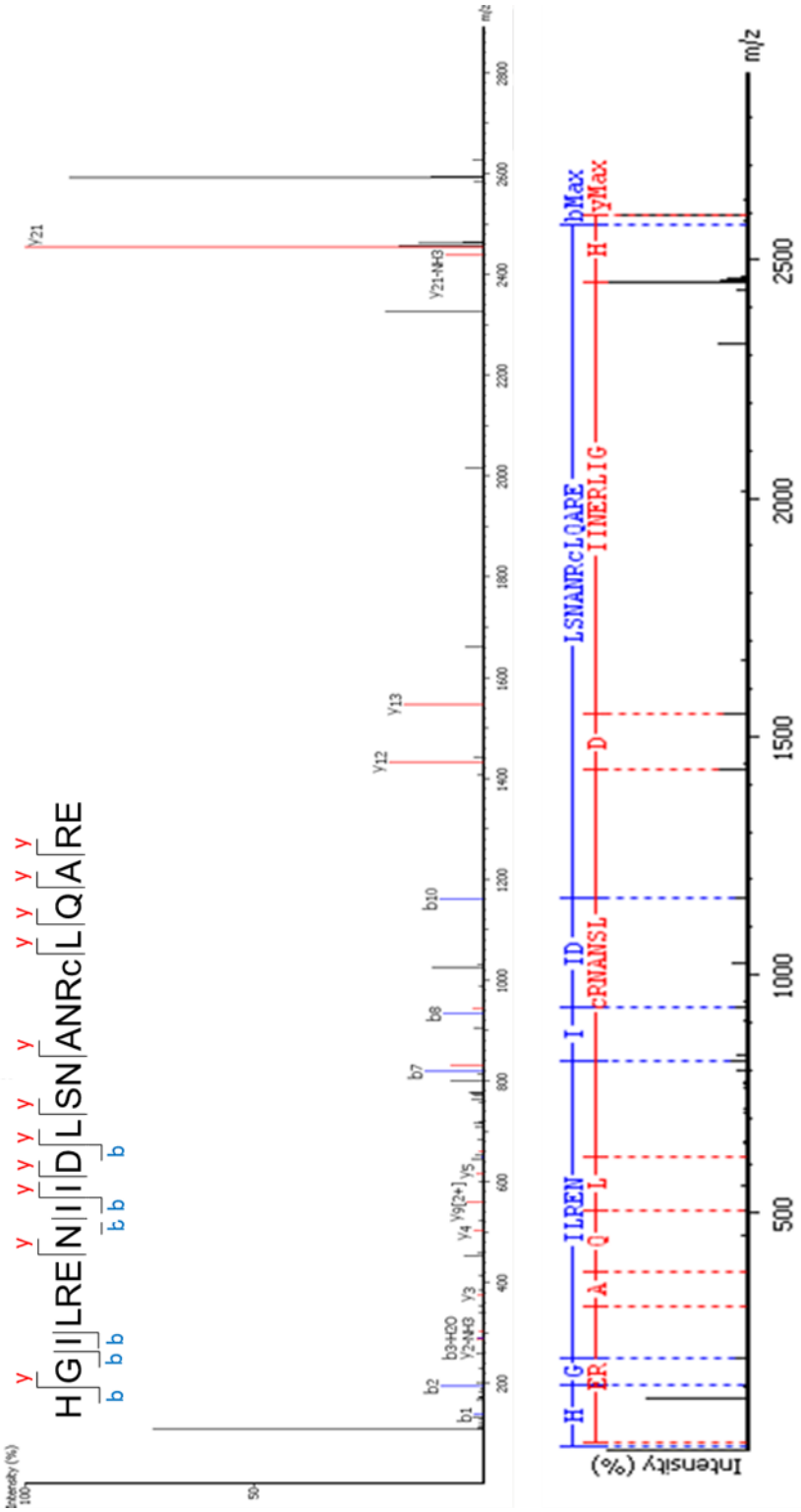


Figure 3.21 *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2592 Da of 18742 Da MUP. Fraction 1 collected from SAX chromatography was digested in-solution using Lys-C as described in hapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.

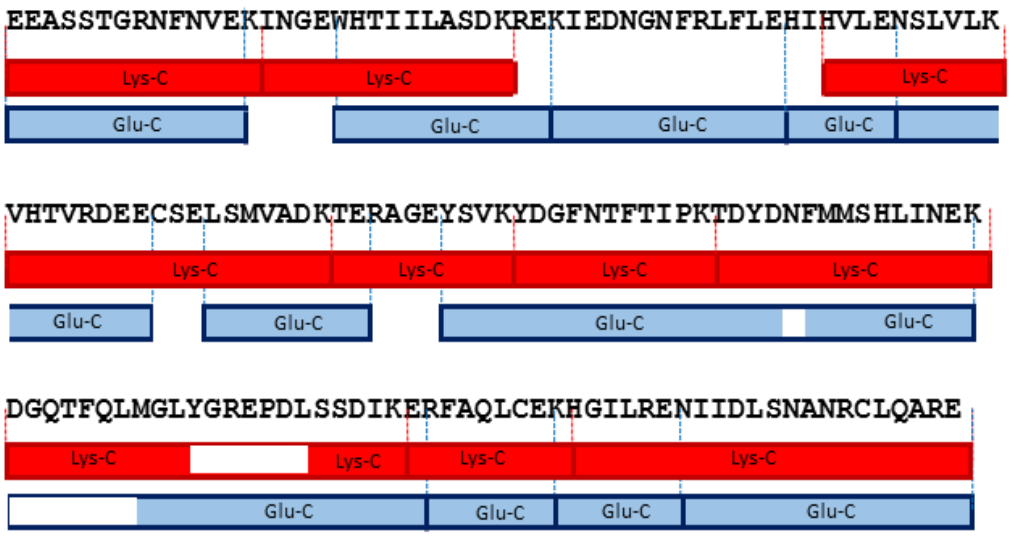


Figure 3.22 Peptide map of the male-specific 18742 Da MUP.
The sequence of the MUP can be confirmed by analysis of the peptides generated separately using two different proteases, Lys-C (in red) and Glu-C (in blue). The solid coloured blocks show sequence confirmation with full peptide sequences. The white blocks highlight unsequenced parts of the peptide.

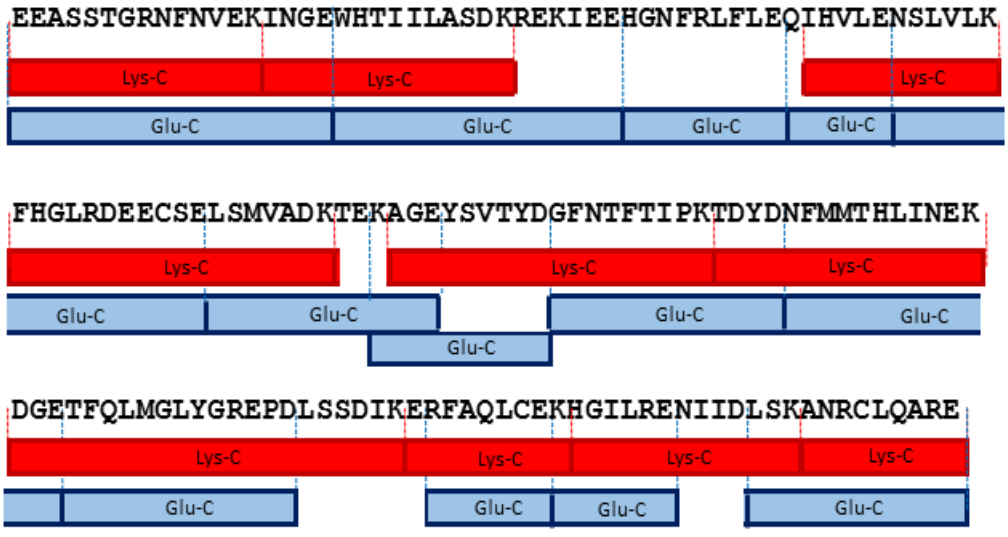


Figure 3.23 Peptide map of the male-specific 18762 Da MUP.
The sequence of the MUP can be confirmed by analysis of the peptides generated separately using two different proteases, Lys-C (in red) and Glu-C (in blue). The solid coloured blocks show sequence confirmation with full peptide sequences.

Figure 3.24 Peptide map of the male-specific 18585 Da MUP.
The sequence of the MUP can be confirmed by analysis of the peptides generated separately using two different proteases, Lys-C (in red) and Glu-C (in blue). The solid coloured blocks show sequence confirmation with full peptide sequences. The white blocks highlight unsequenced parts of the peptide.

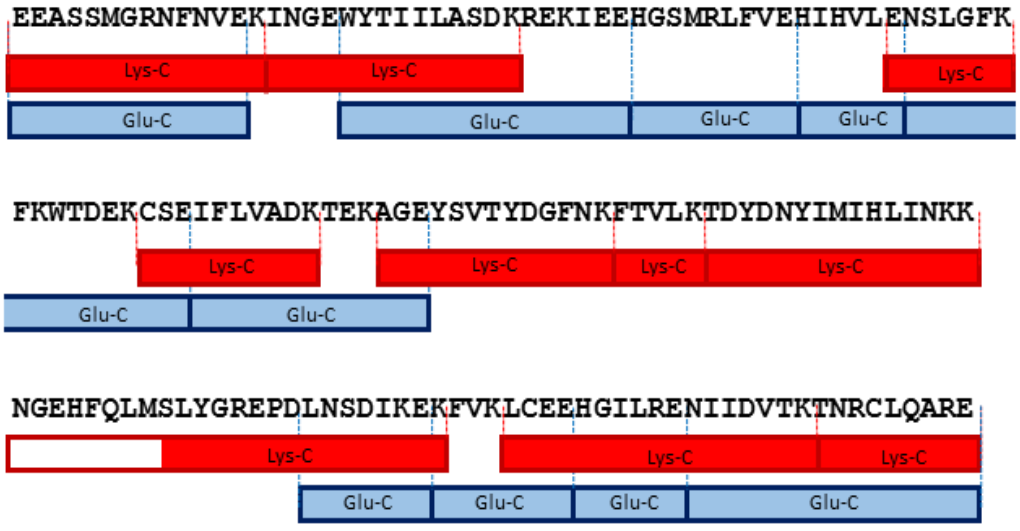


Figure 3.25 Peptide map of the non-sex specific 18918 Da MUP.
The majority of the sequence of the MUP can be confirmed by analysis of the peptides generated separately using two different proteases, Lys-C (in red) and Glu-C (in blue). The solid coloured blocks show sequence confirmation with full peptide sequences. The white blocks highlight unsequenced parts of the peptide.

sequence NGEHFQLM. This sequence was searched for in MS/MS spectra, but no product ions could be identified. For the purpose of aligning *M. spicilegus* MUP sequences, this sequence was used to complete the 18918 Da MUP sequence, due to the relatively high likelihood of this being the correct sequence when considering it results in the correct protein molecular weight (determined by intact protein analysis) and the similarity to the known MUP 6 sequence, which is seen throughout the 18918 Da MUP.

Once the full protein sequences for each *M. spicilegus* MUP were constructed, they were aligned with *M. m. domesticus* MUPs using online sequence alignment tool Clustal Omega (http://www.ebi.ac.uk), and viewed in JalView. Male *M. spicilegus* MUPs were very similar to each other in terms of sequence, and the non sex specific 18918 Da MUP was less homologous to all the male-specific MUPs. Figure 3.26 highlights the sequence differences between each of the *M. spicilegus* MUPs. The similarity of the male-specific MUPs 18742 Da, 18762 Da, and 18585 Da to central *M. m. domesticus* MUPs, particularly male-specific MUP 7, is shown in Figure 3.27, which also confirms the sequence similarity of the *M. spicilegus* 18918 Da MUP found in both sexes to peripheral MUPs.

To further confirm the sequences of the proteins, and to determine which MUP corresponded to each of the protein bands in SDS-PAGE analysis of *M. spicilegus* urine, peptide mass fingerprinting using the MALDI-ToF-MS spectra generated previously from the Lys-C in-gel digestions of the MUP bands (Figures 3.2, 3.6 – 3.8) was undertaken by searching against a custom made database of the *M. spicilegus* MUPs. In the larger, upper male MUP band, peptide masses matching some of those present in all four *M. spicilegus* MUPs were identified. Numerous peptide masses common between the three male specific MUPs were identified, as well as masses unique to the 18742 Da and 18918 Da MUPs. The previously unidentified peptide masses of 1139 Da, 1147 Da, 1302 Da, 1450 Da, 1857 Da and 2038 Da in the spectrum were identified as *M. spicilegus* MUP peptide masses. The presence of peptide masses unique to the 18742 Da and 18918 Da MUPs, and high peptide map coverage of 18585 Da MUP, suggest that this larger protein band corresponds to these three proteins. In the smaller, lower male MUP band, peptide masses matching many of those present in the three male specific *M. spicilegus* MUPs were identified. Again, a number of peptide masses common between these three MUPs were identified, and in this band, peptide masses unique to 18742 Da and 18762 Da MUPs. The previously unidentified masses of 1117 Da, 1139 Da, 1302 Da, 1507 Da, 1857

Da, 1872 Da and 2123 Da were identified as male specific *M. spicilegus* MUP peptide masses. The presence of a number of peptide masses unique to the 18762 Da MUP indicates that this smaller band corresponds to this protein. The identification of a peptide mass unique to the 18742 Da protein in this spectrum is likely to be due the fact a small amount of the larger, upper band was also removed when excising this band from the gel for digestion. In the female MUP band, peptide masses matching to the 18918 Da MUP peptides were identified – all these masses were unique to the 18918 Da MUP, and there was no evidence of peptides from any of the male-specific MUPs. In this spectrum, the previously unidentified masses of 1450 Da and 2038 Da were identified as 18918 Da MUP peptide masses.

No complete peptide mass fingerprints were obtained for any of the four *M. spicilegus* MUPs; the L4 peptides of each of the male specific MUPs were not identified in any spectra. In LC-MS/MS analysis of the Lys-C digestions of these proteins, this peptide was not identified or sequenced (in each protein, it was successfully sequenced in the analysis of Glu-C digestions). The L4 peptide of central MUP 7 (2840.5 Da) has previously been determined to ionise very poorly (Dr. S. Armstrong, Ph.D thesis), and the sequence similarity of this peptide to the male specific *M. spicilegus* MUP L4 peptides indicate that poor ionisation is the reason this peptide isn't observed in the MALD-ToF-MS spectra. The 18585 Da L5 peptide, not identified in the spectra in Figures 3.28 – 3.29, was identified in the analysis of the Lys-C digestion of other 'upper' male MUP bands (Supplementary Material). Four of the 18918 Da peptide masses were not identified in the MALDI-ToF-MS spectrum of the female band, likely to be due to the fact that this MUP is only present in very small amounts in urine samples. The peptide mapping of these four *M. spicilegus* MUPs, plus the identification of a number of previously unidentified peptide masses in male and female MALDI-ToF-MS spectra of digested *M. spicilegus* MUP bands, further supports that the LC-MS/MS analysis and *de novo* sequencing have enabled the correct sequence determination of all four *M. spicilegus* proteins.

Figure 3.26 Alignment of the final *M. spicilegus* MUP sequences.
The four MUP sequences were aligned using Clustal Omega (www.ebi.ac.uk) and the multiple alignment was viewed using Jalview. Residues were coloured using the Clustal option to highlight sequence differences.

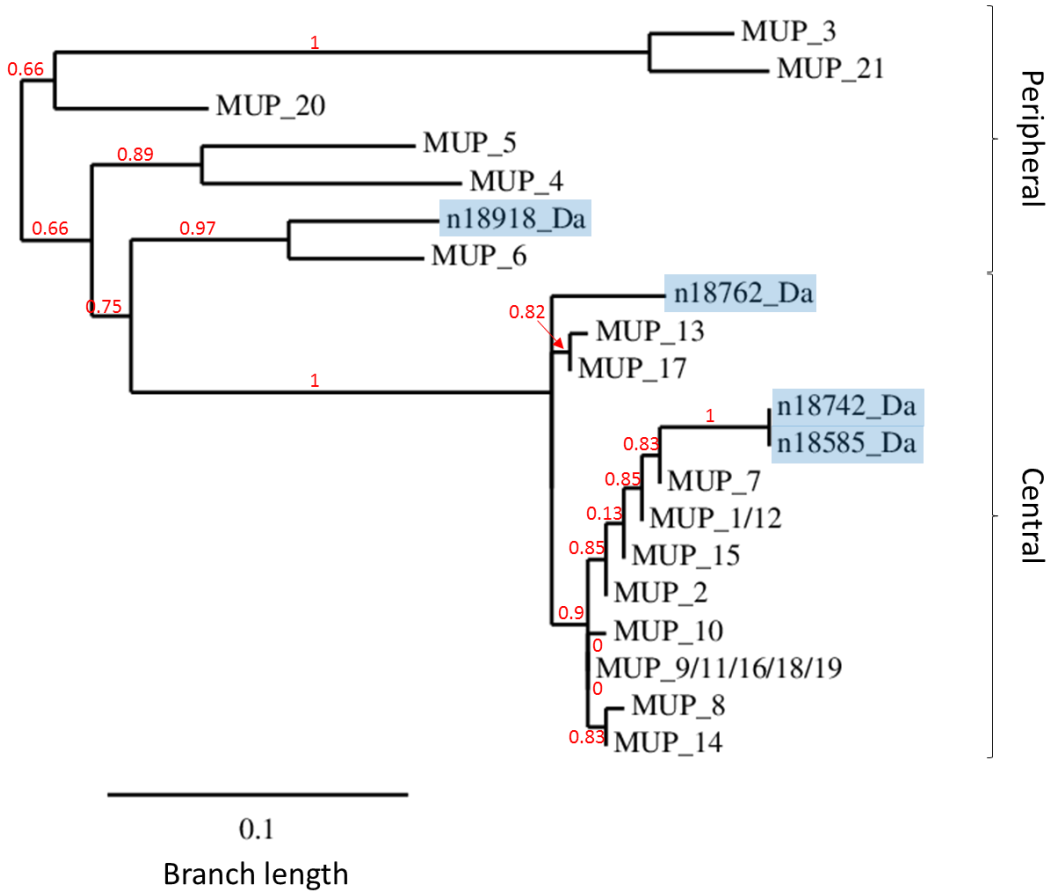


Figure 3.27 Phylogentic tree of *M. spicilegus* and *M. m. domesticus* MUPs. Phylogenetic analysis was performed using the 'One Click' mode at www.phylogeny.fr (Dereeper *et al.* 2008). Sequences were aligned using MUSCLE (v3.8.31), configured for highest accuracy. Gblocks (v0.91b) was used to remove any ambiguous regions. The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT). Graphical representation of the tree was performed with TreeDyn (v198.3). Bootstrap values are labelled in red, peripheral and central *M. m. domesticus* MUPs are identified, and *M. spicilegus* MUPs are highlighted in blue.

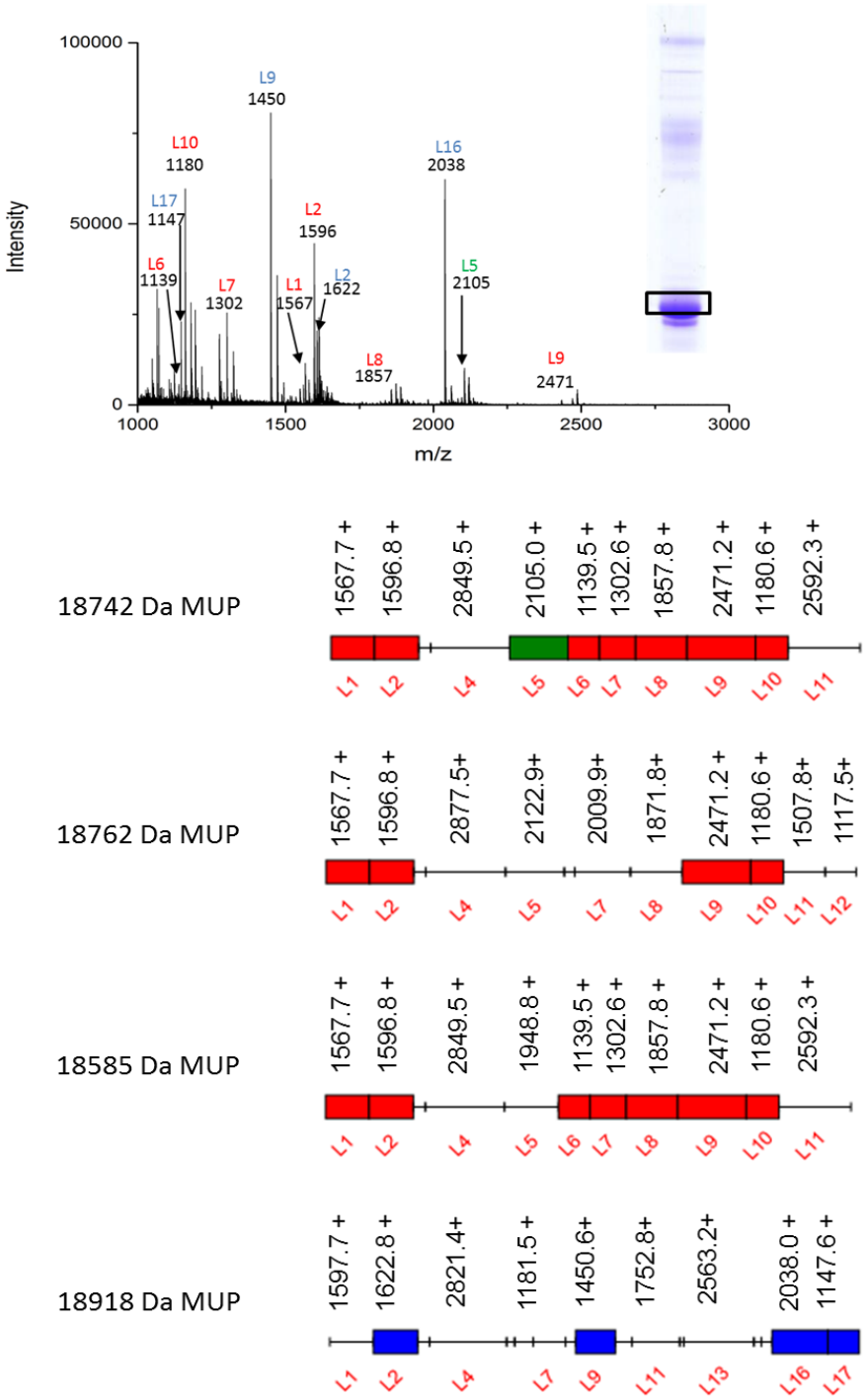


Figure 3.28 *M. spicilegus* MUP peptide mass fingerprinting of a male *M. spicilegus* urine sample.
The MALDI-ToF-MS spectrum generated from a Lys-C in-gel digestion of SDS-PAGE gel pieces (Figures 3.2 and 3.6) was searched against a custom made *M. spicilegus* MUPs database using Mascot. The digested protein band from male urine showed peptides that matched peptides from all *M. spicilegus* MUPs, the peptide maps of which are shown below the spectrum. In the peptide maps, the coloured peptides are those masses observed in the spectrum. The red peptides are those with masses shared between two or more of the *M. spicilegus* MUPs, and peptide masses unique to each MUP are coloured differently to one another – peptide masses unique to 18742 Da are coloured green; to 18762 Da, purple; to 18918 Da; blue. No unique peptides from the 18585 Da MUP were observed.

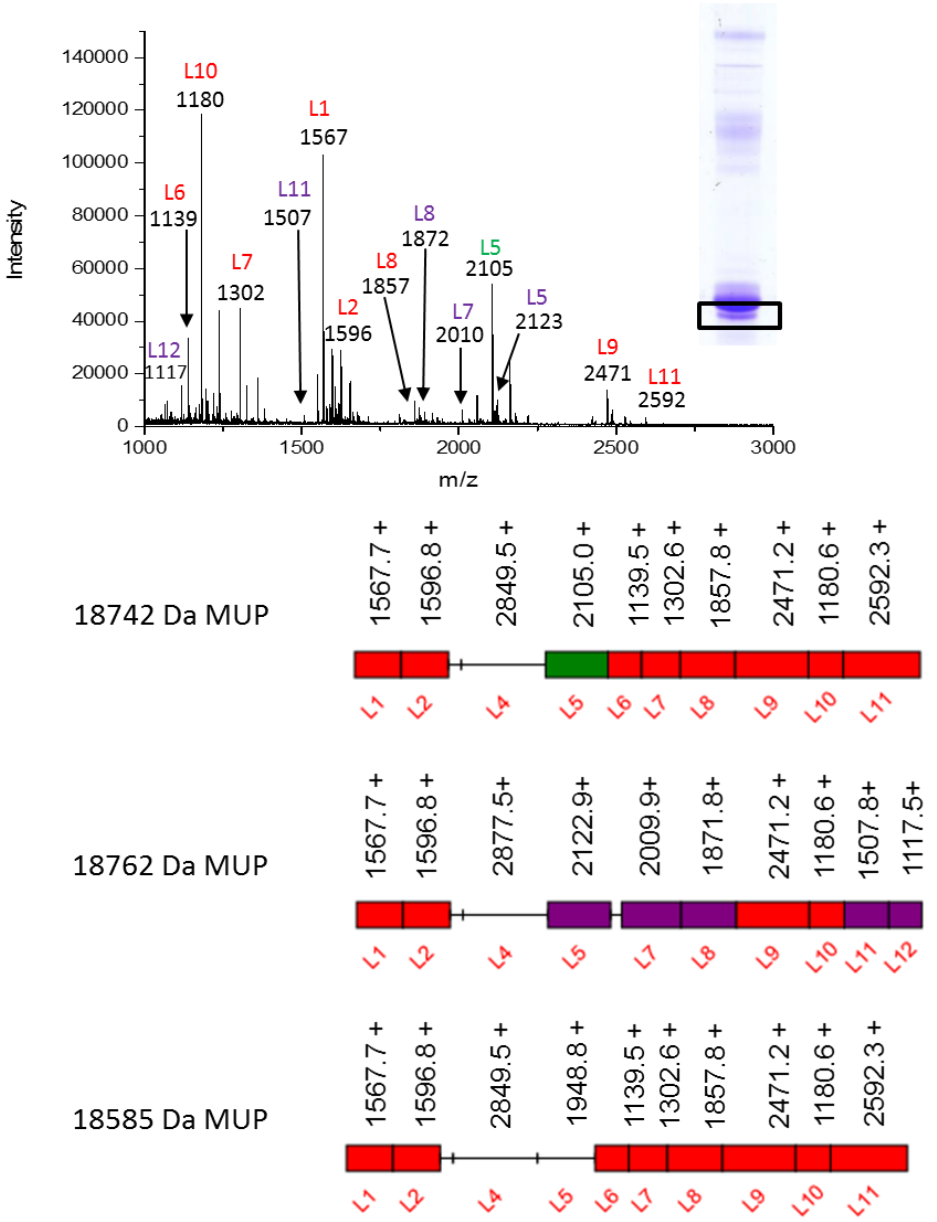


Figure 3.29 *M. spicilegus* MUP peptide mass fingerprinting of a male *M. spicilegus* urine sample.
The MALDI-ToF-MS spectrum generated from a Lys-C in-gel digestion of SDS-PAGE gel pieces (Figures 3.2 and 3.7) was searched against a custom made *M. spicilegus* MUPs database using Mascot. The digested protein band from male urine showed peptides that matched peptides from *M. spicilegus* MUPs, the peptide maps of which are shown below the spectrum. In the peptide maps, the coloured peptides are those masses observed in the spectrum. The red peptides are those with masses shared between two or more of the *M. spicilegus* MUPs, and peptide masses unique to each MUP are coloured differently to one another – peptide masses unique to 18742 Da are coloured green; to 18762 Da, purple. No unique peptides from the 18585 Da or 18918 Da MUPs were observed.

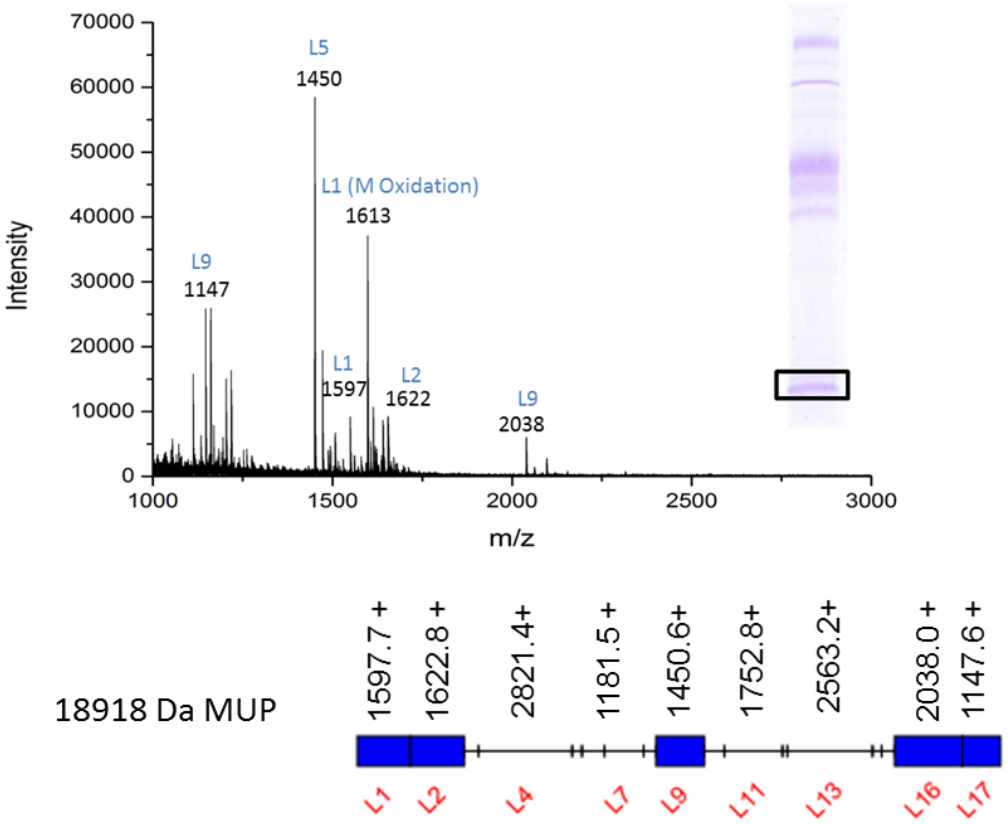


Figure 3.30 *M. spicilegus* MUP peptide mass fingerprinting of a female *M. spicilegus* urine sample.
The MALDI-ToF-MS spectrum generated from a Lys-C in-gel digestion of SDS-PAGE gel pieces (Figures 3.2 and 3.8) was searched against a custom made *M. spicilegus* MUPs database using Mascot. The digested protein band from female urine showed peptides that matched peptides from one *M. spicilegus* MUP, the peptide map of which is shown below the spectrum. In the peptide maps, the blue peptides are those masses observed in the spectrum. No peptides from the 18742 Da, 18762 Da or 18585 Da MUPs were observed.

3.2.7 Investigating the roles of *M. spicilegus* MUPs in their behaviour

The MUP pattern in *M. spicilegus*, though simpler than in *M. m. domesticus*, is variable between individual males, suggesting that the MUPs do have a degree of polymorphism. Whilst inbred laboratory strains of the house mouse exhibit very similar MUP expression patterns in their urine, which is expected since inbreeding results in genetically identical mice, this random-bred strain of *M. spicilegus* (particularly males) exhibit variable MUP expression. When the behavioural studies previously conducted with this species are considered, it is possible that the MUPs play an important role in either reproductive behaviour, social behaviour, or both. The fact that female *M. spicilegus* are very aggressive towards each other, yet their MUP expression patterns are extremely simple and do not vary between individuals, may point to the male-specific MUPs being involved in reproductive behavioural characteristics. As a result, a more in-depth experiment that includes the recovery of mouse urine and assessment of MUP expression during different reproductive states and social situations would give an insight into the function of *M. spicilegus* MUPs.

Three different experiments were set up at the Leahurst campus, where the *M. spicilegus* mice were based, to determine whether contact between males and females, breeding between males and females and contact between males caused differences in their MUP expression output in the aim of determining a possible role of MUPs in their behaviour. Prior to the experiments, male mice were singly housed and female mice were housed with their female siblings. All had no prior contact with the opposite sex as adults.

Experiment 1 – Contact between male and female mice

In the first experiment, five non-sibling male/female pairs were placed in large split cages, where a mesh barrier was placed between them through which they could contact each other, but not gain full access (Figure 3.31). The male/female pairs were housed in this setting for 23 days, with their urine being sampled using the recovery method on the day prior to them coming into contact, one day after they began contact, ten days after they began contact and on the final day. The aim of this experiment was to determine whether the mice, particularly males, exhibited different MUP expression patterns upon initial contact with a mouse of the opposite sex that they had not encountered before, and whether MUP expression patterns changed with prolonged contact with the opposite sex.

All urine samples were frozen upon collection until further analysis. All samples recovered from male mice were analysed for the presence and relative abundance of MUPs using SDS-PAGE, and samples were then analysed using ESI-MS to provide an insight into the relevant amounts of each MUP expressed at each stage of the experiment. If there was sample remaining, protein and creatinine assays were carried out to confirm the changing concentration of protein in the urine. All samples recovered from female mice were initially subjected to SDS-PAGE analysis, and upon initial ESI-MS analyses of female urine, it was determined that the amount of MUP in the majority of female samples was too low to be detected.

All five paired female mice showed no difference in their urinary protein output during the course of the contact experiment, as assessed by SDS-PAGE (Figure 3.32). All show a very low intensity band on the SDS-PAGE gels at approximately 20 kDa, which from previous analyses is the 18918 Da MUP. In male samples, MUP expression patterns were similar between individuals prior to contact with their female pair, but differed between individuals in terms of the amount of MUP expressed. The male-specific 18742 Da and 18762 Da MUPs were identified in the urine of most males prior to female contact, with the male-specific 18585 Da and non-sex specific 18918 Da MUPs in two of the male samples obtained prior to contact. The intensity of the MUP bands in SDS-PAGE gave an indication of the amount of MUP present in the samples, confirmed by the calculation of protein:creatinine ratios (Figure 3.33 – 3.37). In four of the five males, MUP output increased upon contact with the female (after one day of contact) and had increased further after ten days of contact. One male, however, expressed less MUP when urine was collected on the final day of the experiment (Figure 3.34). ESI-MS analysis enabled the identification of differing MUP expression patterns in different samples, quantifying MUPs relative to the others in the sample. Of particular interest is the significant increase of the male-specific 18762 Da MUP upon continuing contact with a female – prior to contact with females, and in all other intact mass analyses of male urine until this point, the most abundant MUP relative to the others has always been the male-specific 18742 Da MUP (usually the 18762 Da MUP is only approximately 20 – 50% of the intensity of the 18742 Da MUP in ESI-MS spectra). As the time in contact with females increases, the male expression of the 18762 Da MUP increased (Figures 3.33 – 3.37). In ESI-MS data analysis, the MaxENT 1 function (based on maximum entropy) measures the relatively quantitative intensities of the sample components by summing the peak intensities from the original multiply-charged spectra, and this is represented by the area under the molecular weight peak in the processed spectra (Cottrell and Green

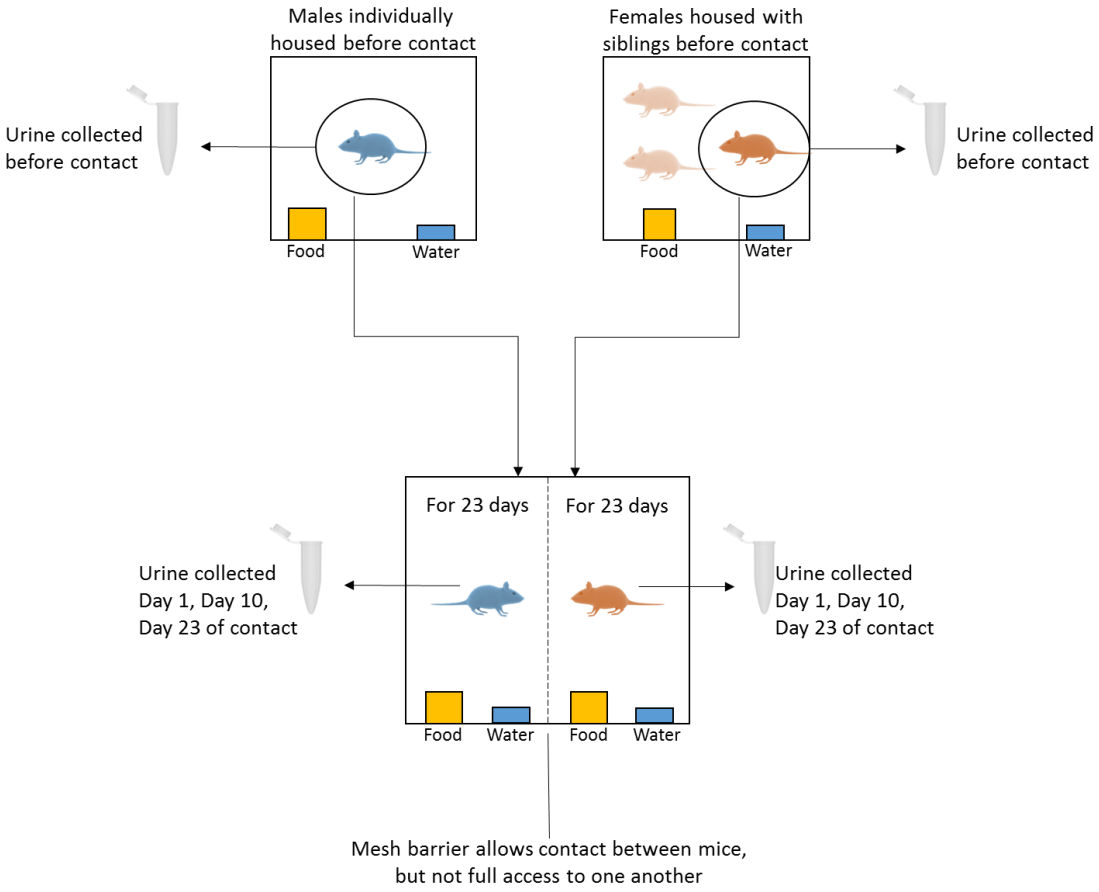


Figure 3.31 Set-up of Experiment 1.
Set-up of Experiment 1 – Contact between male and female mice. Urine samples were taken from a singly-housed male and an unrelated group-housed female before the pair were housed in a split cage, separated by a mesh barrier which allowed contact between the male and female, but not full physical access to one another. Male and female pairs were housed in the split cages for 23 days, with urine collected from mice on days 1, 10 and 23 of contact.

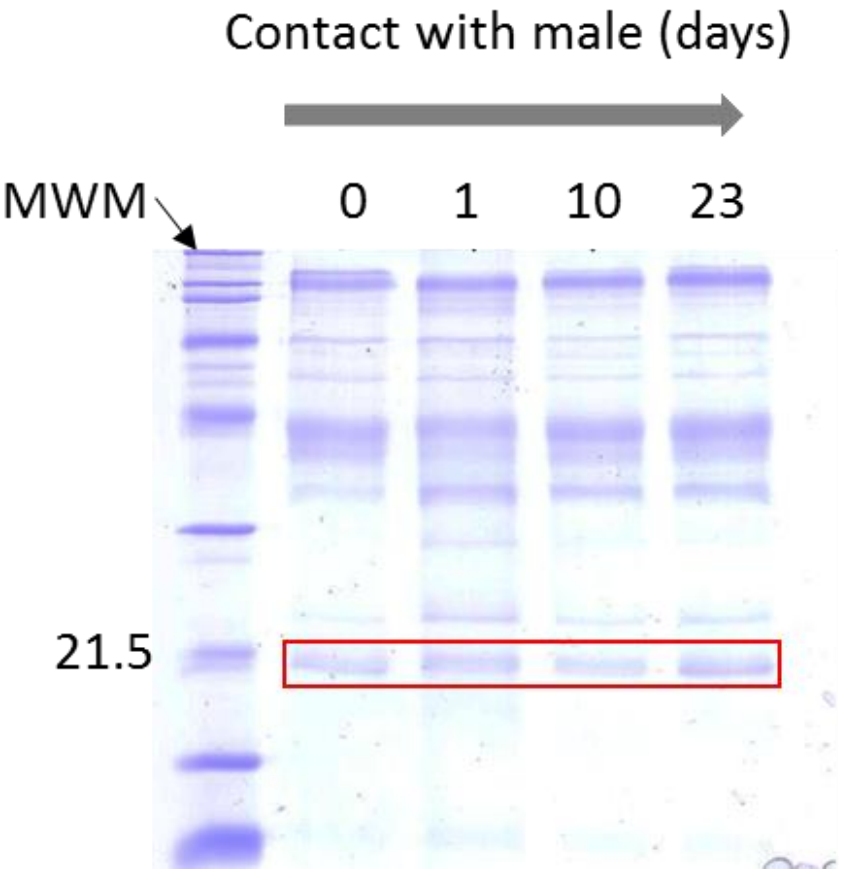


Figure 3.32 SDS-PAGE analysis of female *M. spicilegus* urine.
Urine (5 µl) from a female *M. spicilegus,* taken prior to contact with the male, one day after contact began, after ten days of contact and on the final day of contact, was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. The very low intensity MUP bands at approximately 20 kDa are highlighted in red.

Figure 3.33 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and on the final day of contact. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

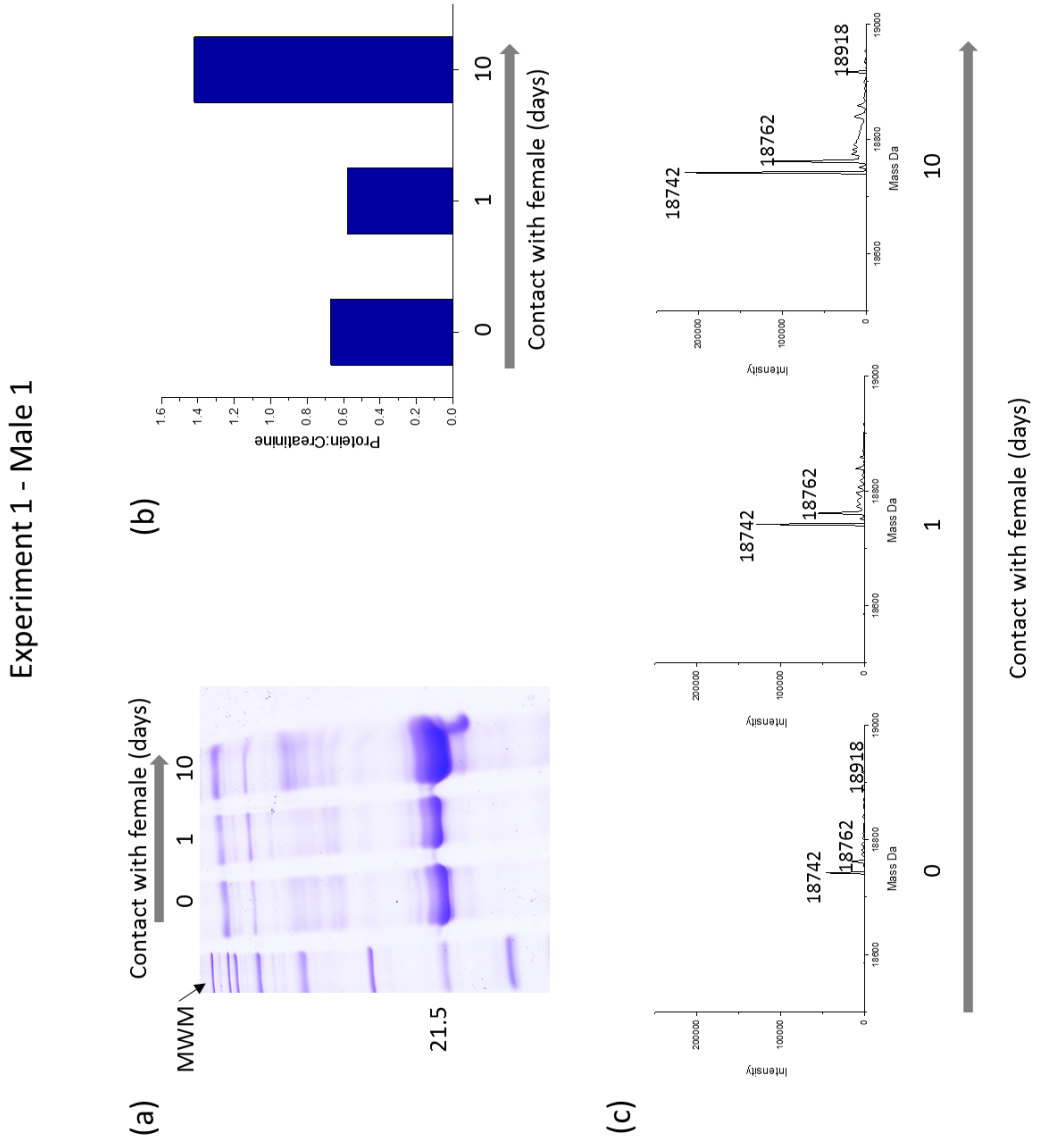

Experiment 1 - Male 1
(a)
Contact with female (days)
0
1
10
MWM
21.5
(b)
Protein:Creatinine
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
1
10
Contact with female (days)
(c)
18742
18762
18918
Intensity
Mass Da
0
1
10
Contact with female (days)

Figure 3.34 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and on the final day of contact. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 3.1, Waters).

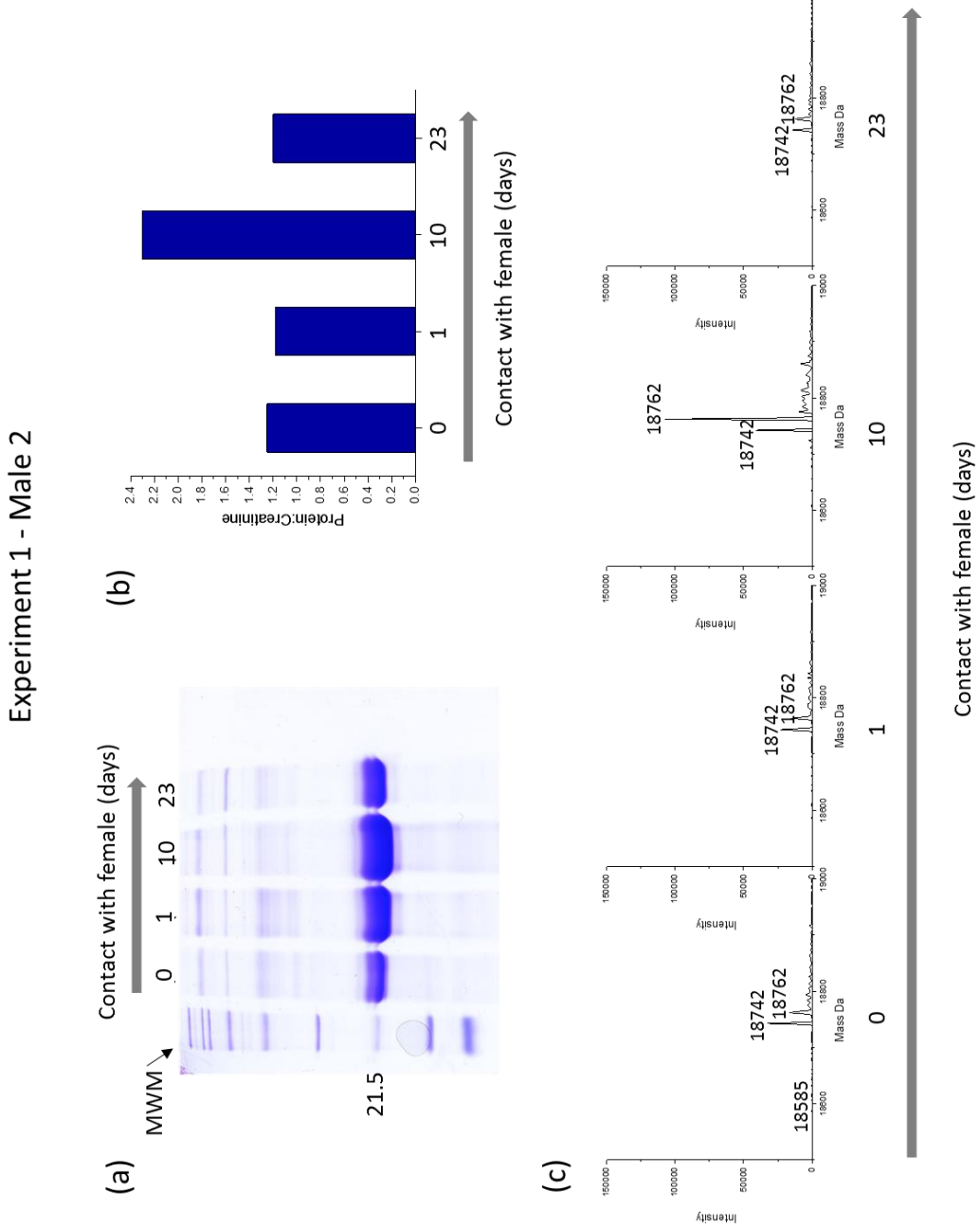

Experiment 1 - Male 2
(a)
MWM
Contact with female (days)
0
1
10
23
21.5
(b)
Protein:Creatinine
Contact with female (days)
(c)
18585
18742
18762
Intensity
Mass Da
Contact with female (days)

Figure 3.35 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and on the final day of contact. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

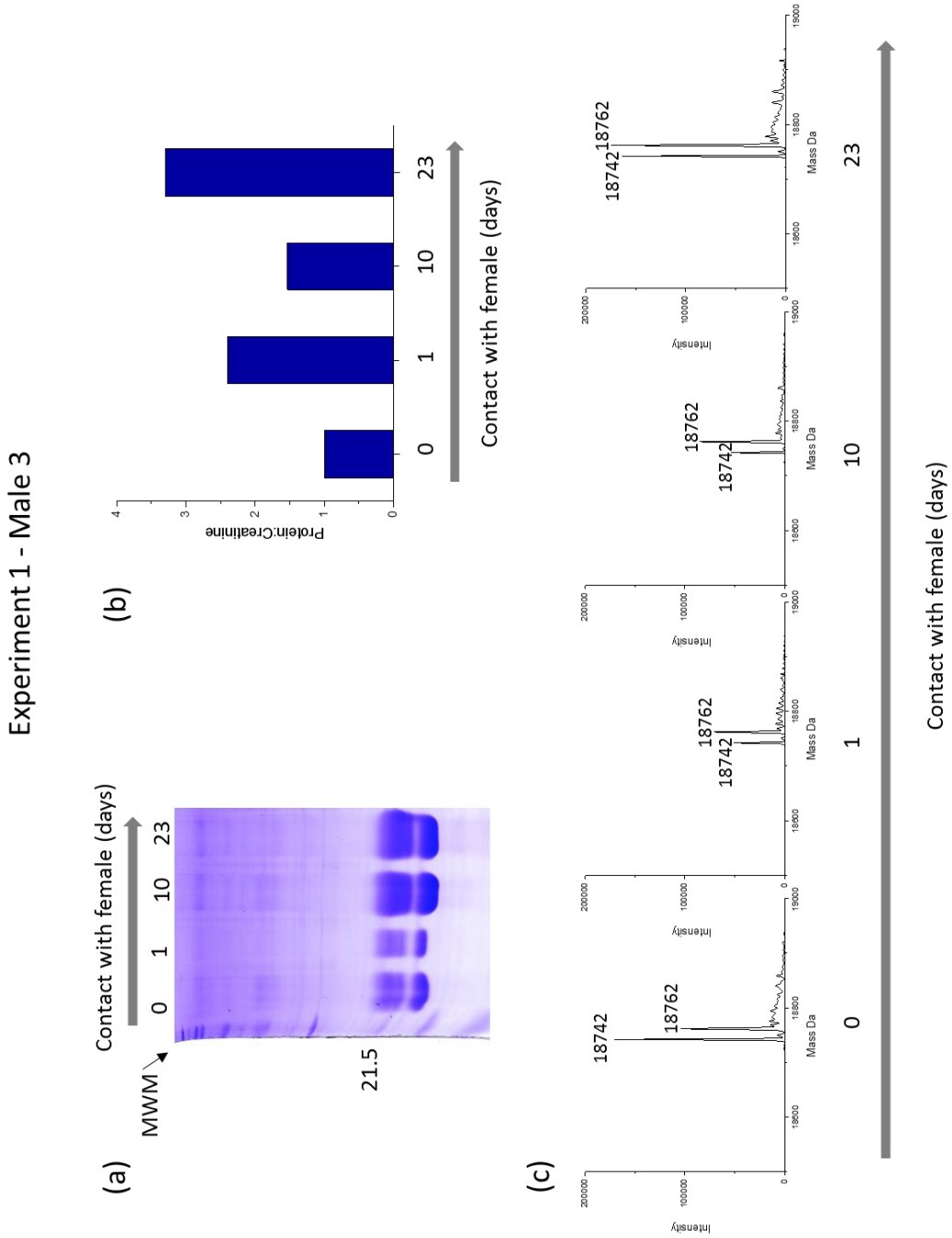

Experiment 1 - Male 3
(a)
Contact with female (days)
0
1
10
23
MWM
21.5
(b)
Protein:Creatinine
0
1
2
3
4
0
1
10
23
Contact with female (days)
(c)
18742
18762
Intensity
Mass Da
0
1
10
23
Contact with female (days)

Figure 3.36 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and on the final day of contact. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution (NS = not enough sample for protein and creatinine assays). (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

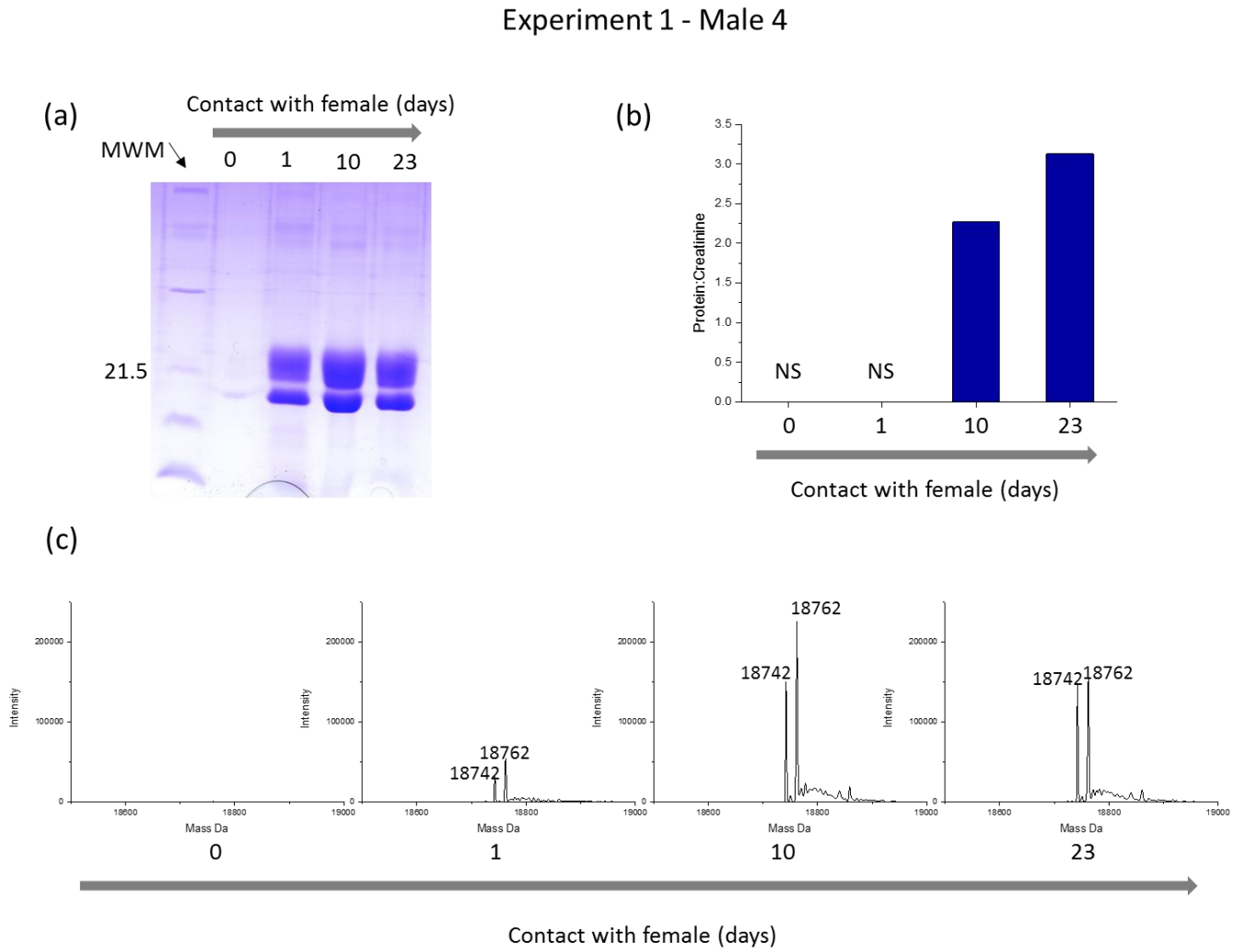

Experiment 1 - Male 4
(a)
Contact with female (days)
MWM
0
1
10
23
21.5
(b)
Protein:Crealtinine
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
NS
NS
0
1
10
23
Contact with female (days)
(c)
Intensity
Mass Da
18742
18762
0
1
10
23
Contact with female (days)

Figure 3.37 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and on the final day of contact. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

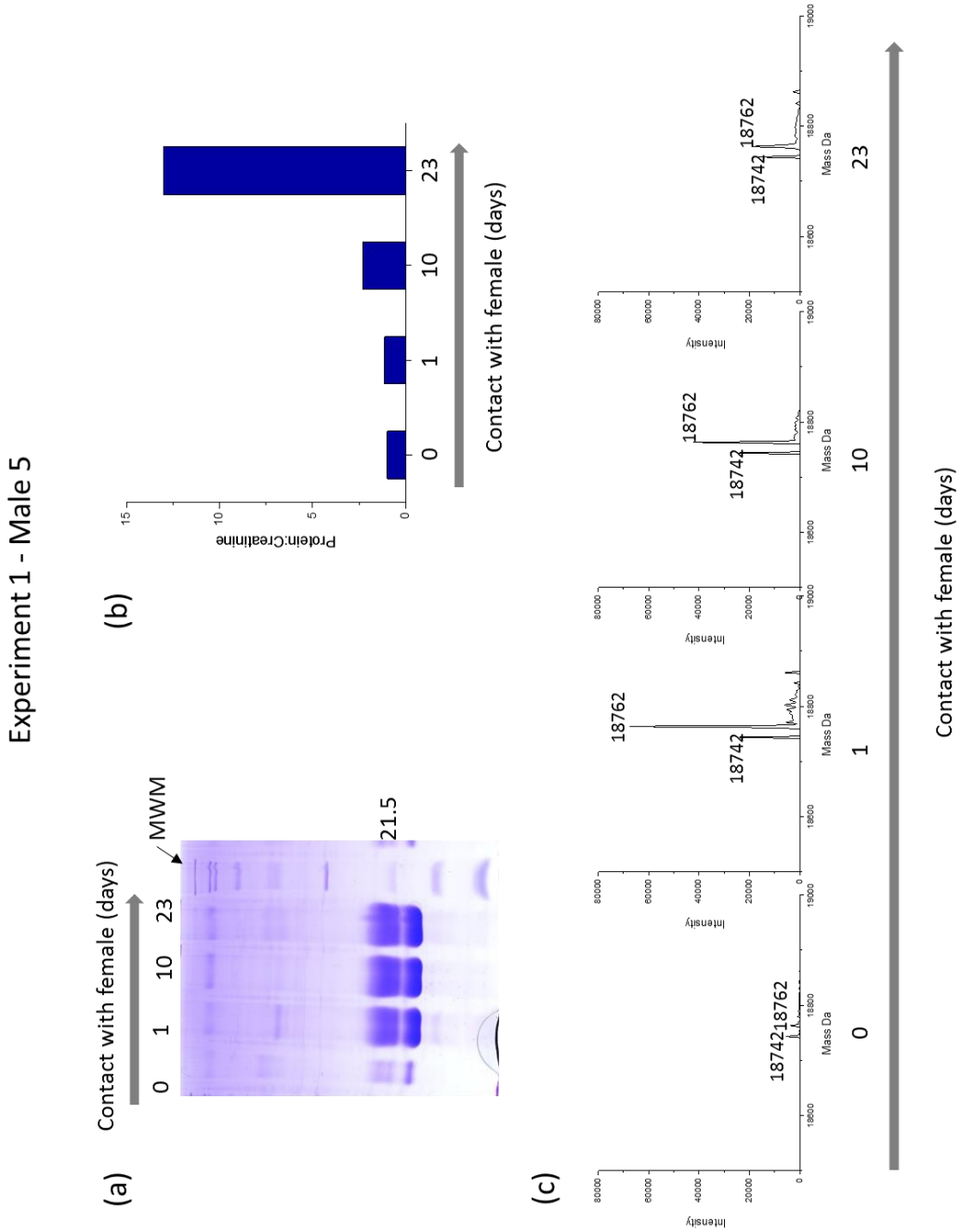
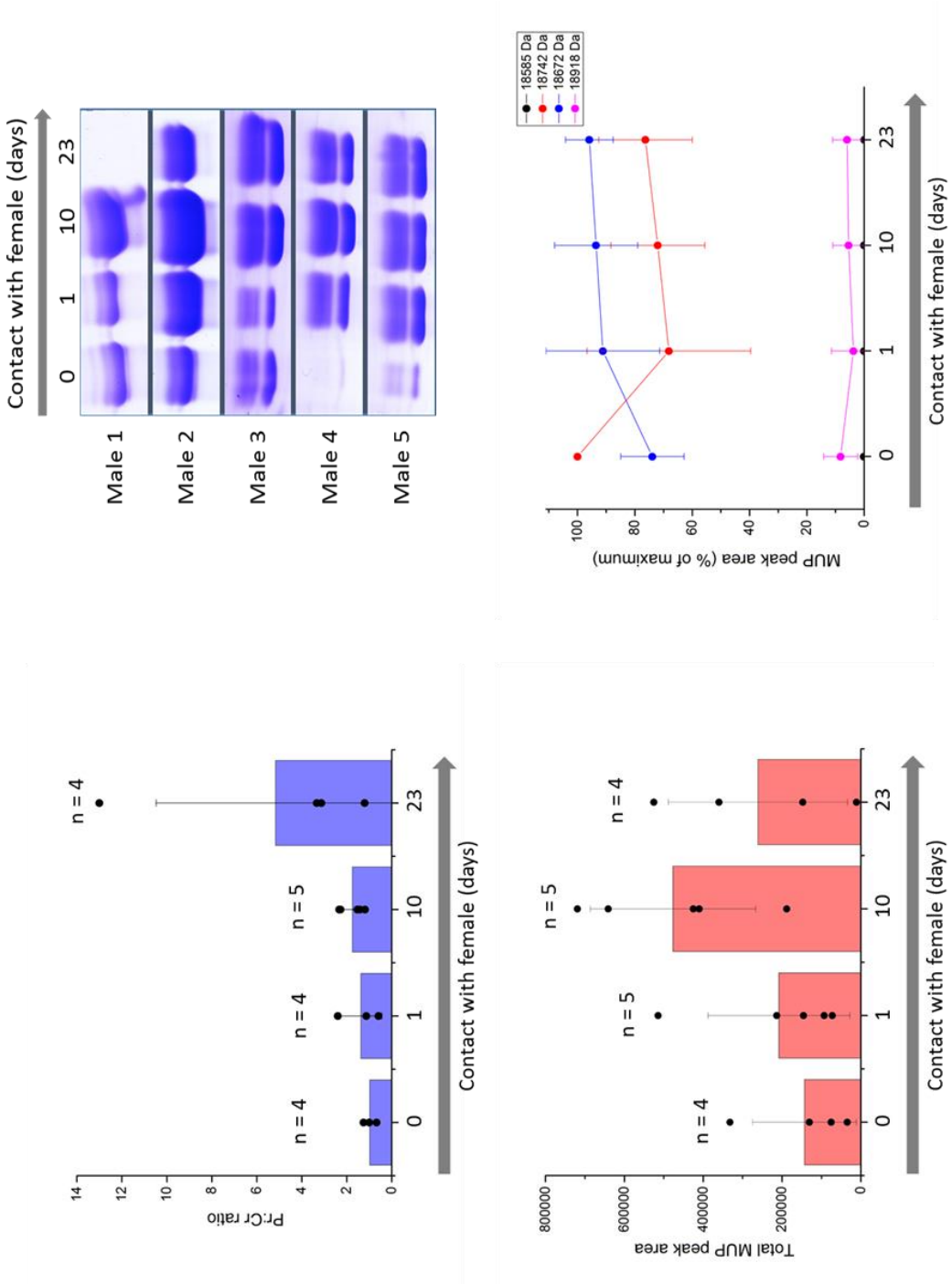

Experiment 1 - Male 5
(a)
Contact with female (days)
0
1
10
23
MWM
21.5
(b)
Protein:Creatinine
0
1
10
23
Contact with female (days)
(c)
Intensity
Mass Da
18742
18762
0
1
10
23
Contact with female (days)

Figure 3.38 Summary of the analyses of *M. spicilegus* urine from male mice in Experiment 1.
Top left: The average protein:creatinine ratios calculated for the urine samples from each of the five males at each sample collection point (error bars ± SD). Data points for each sample are shown. Top right: The ~20 kDa region (MUP bands) of the SDS-PAGE analysis of the urine from each of the five males. Bottom left: The average total peak area of all *M. spicilegus* MUP peaks in ESI-MS analysis (error bars ± SD). MUP peak areas (providing a relative quantification) observed in each of the ESI-MS spectra were calculated using MaxENT software and summed for each sample. Data points for each sample are shown. Bottom right: The average peak areas of each *M. spicilegus* MUP peak in the ESI-MS analysis (error bars ± SD). The peak area of each MUP peak in each spectrum were calculated using MaxENT software. For each male urine sample, the MUP peak areas were expressed as a percentage of the maximum MUP peak area.


Contact with female (days)
0
1
10
23
Male 1
Male 2
Male 3
Male 4
Male 5
18585 Da
18742 Da
18672 Da
18918 Da
MUP peak area (% of maximum)
Contact with female (days)
Total MUP peak area
n = 4
n = 5
Contact with female (days)
Pr:Cr ratio
n = 4
n = 4
n = 5
n = 4
Contact with female (days)

1998). Figure 3.38 shows that the total MUP peak area increases in each male sample as time in contact with a female increases up until day 10 of contact, yet decreases slightly by day 23 of contact. Upon contact with females, the amount of the 18762 Da MUP (relative to the other MUPs in the sample) increases in all male urine samples, and remains the most abundant MUP throughout the experiment (Figure 3.38). The increase in the amount of 18762 Da MUP expressed with female contact compared to the amount expressed before contact was not significant (Welch t-test, $p = 0.99$), and the decrease in the 18742 Da MUP expressed with female contact compared to before contact was also not significant (Welch t-test, $p = 0.32$).

Experiment 2 – Breeding between male and female mice

In the second experiment, five non-sibling male/female pairs were placed in large split cages, where initially a mesh barrier was placed between them through which they could contact each other (Figure 3.39). The male/female pairs were housed in this setting for ten days, with urine samples collected from each animal prior to contact, after one day of being in contact and after ten days of contact. On the tenth day, the contact barrier was removed to allow the pairs to live together, giving them the opportunity to interact and mate, for a further seven days, after which a final urine sample was recovered from each animal. The aim of this experiment was to determine whether the mice, particularly males, exhibited altered MUP expression patterns upon complete physical contact and possible mating with the opposite sex, compared to through-barrier contact. Out of the five male/female pairs, three of them reproduced.

All urine samples were frozen upon collection until required for analysis. All samples recovered from male mice were analysed for the presence and relative abundance of MUPs using SDS-PAGE. Samples were then analysed using ESI-MS to provide an insight into the relevant amounts of each MUP expressed at each stage of the experiment. If there was sample remaining, protein and creatinine assays were carried out to confirm the changing concentration of protein in the urine. All samples recovered from female mice were initially subjected to SDS-PAGE analysis, and upon initial ESI-MS analyses of female urine, it was determined that the amount of MUP in the majority of female samples was too low to be detected.

SDS-PAGE analysis showed that all five paired female mice showed no difference in their urinary protein output during the course of the contact experiment. The majority of all collected female urine samples show a very low intensity band on the SDS-

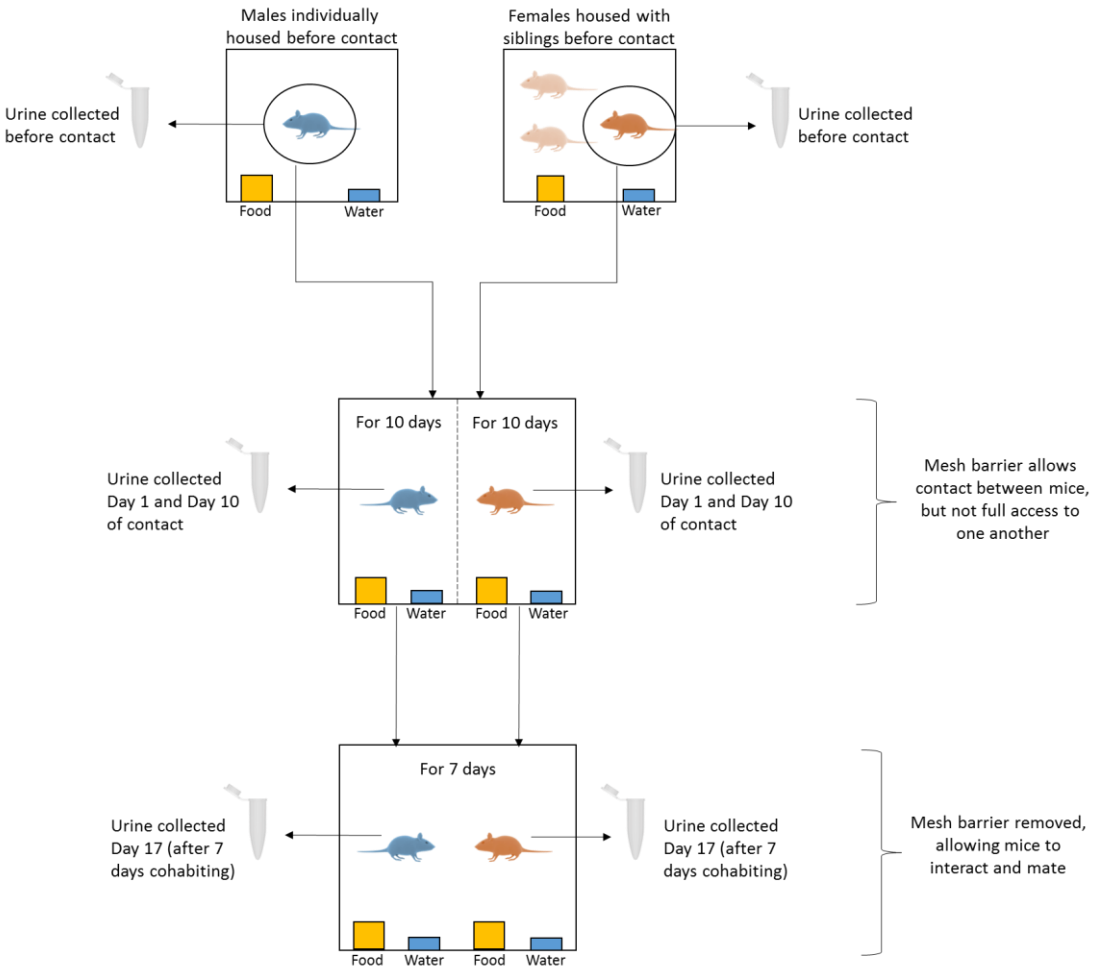


Figure 3.39 Set-up of Experiment 2.
Set-up of Experiment 2 – Breeding between male and female mice. Urine samples were taken from a singly-housed male and an unrelated group-housed female before the pair were housed in a split cage, separated by a mesh barrier which allowed contact between the male and female, but not full physical access to one another. Male and female pairs were housed in the split cages for 10 days, with urine collected from mice on days 1 and 10 of contact. On day 10, the mesh barrier was removed, and male/female pairs lived together, able to interact and mate, for a further 7 days. A final urine sample was taken from each animal on the final day of the experiment (Day 17).

PAGE gels at approximately 20 kDa, which from previous analyses is the 18918 Da MUP. In male samples, MUP expression patterns were variable between individuals prior to contact with their female pair as well as differing between individuals in terms of the amount of MUP expressed. The male-specific 18742 Da and 18762 Da MUPs were identified in the urine of all males prior to female contact, with the non sex specific 18918 Da MUPs in four of the male samples obtained prior to contact. The intensity of the MUP bands in SDS-PAGE gave an indication of the amount of MUP present in the samples, confirmed by the calculation of protein:creatinine ratios (Figures 3.40 – 3.44). In three of the males, urinary MUP output increased immediately upon contact with the female (after one day of contact) and all five males increased MUP output after ten days of contact. All males expressed less MUP when urine was collected on the final day of the experiment, after being put together to interact and mate, with three out of the five males expressing substantially less MUP. Protein and creatinine assays confirmed that protein was present in these urine samples, indicating that MUP expression had been significantly reduced (Figures 3.40 – 3.44). Again, of particular interest is the significant increase of the male-specific 18762 Da MUP upon continuing contact with a female – this was observed in the samples from all five males, and of further interest is the fact that this MUP was reduced in expression after complete contact/breeding. Again, as the time in contact with females increases, the male expression of the 18762 Da MUP increases (Figures 3.40 – 3.44), and after complete contact/breeding, reduces again to amounts similar to those seen in the urine taken prior to female contact. Figure 3.45 confirms that the total MUP peak area increases in each male sample as time in contact with a female increases up until day 10 of contact, and decreases after complete contact/breeding. Upon contact with females, the amount of the 18762 Da MUP (relative to the other MUPs in the sample) increases in all male urine samples, and remains the most abundant MUP throughout contact, with the amount of the 18742 Da MUP decreasing. After complete contact/breeding, the expression of these two MUPs return to that identified in the urine prior to female contact (Figure 3.45).

Figure 3.40 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and after being put together for breeding. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

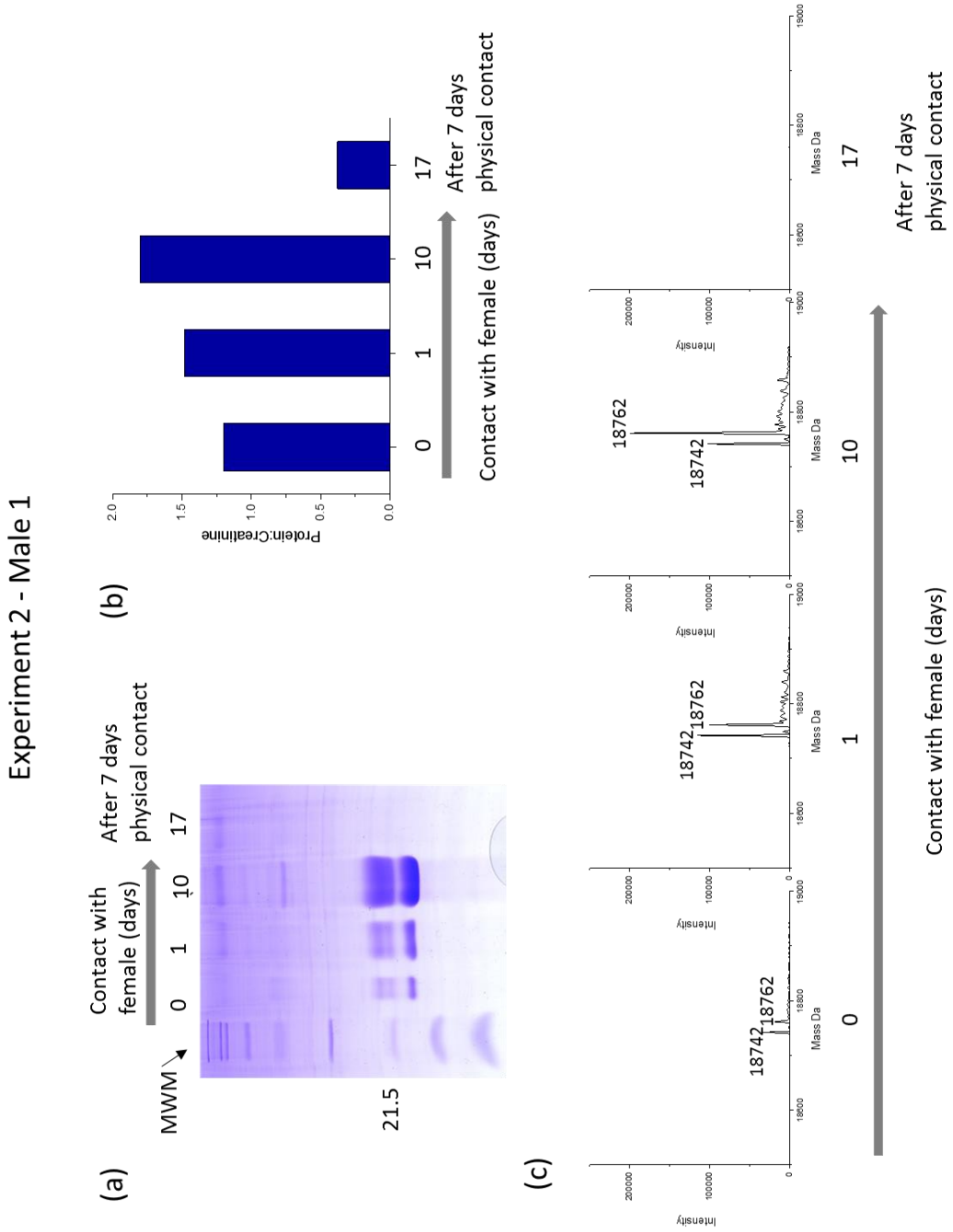

Experiment 2 - Male 1
(a)
MWM
Contact with female (days)
0
1
10
17
After 7 days physical contact
21.5
(b)
Protein:Creatinine
2.0
1.5
1.0
0.5
0.0
0
1
10
17
Contact with female (days)
After 7 days physical contact
(c)
Intensity
Mass Da
18742
18762
0
1
10
17
Contact with female (days)
After 7 days physical contact

Figure 3.41 SDS-PAGE and ESI-MS analyses of male *Mus spicilegus* urine.
Urine (5 µl) from a male *Mus spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and after being put together for breeding. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

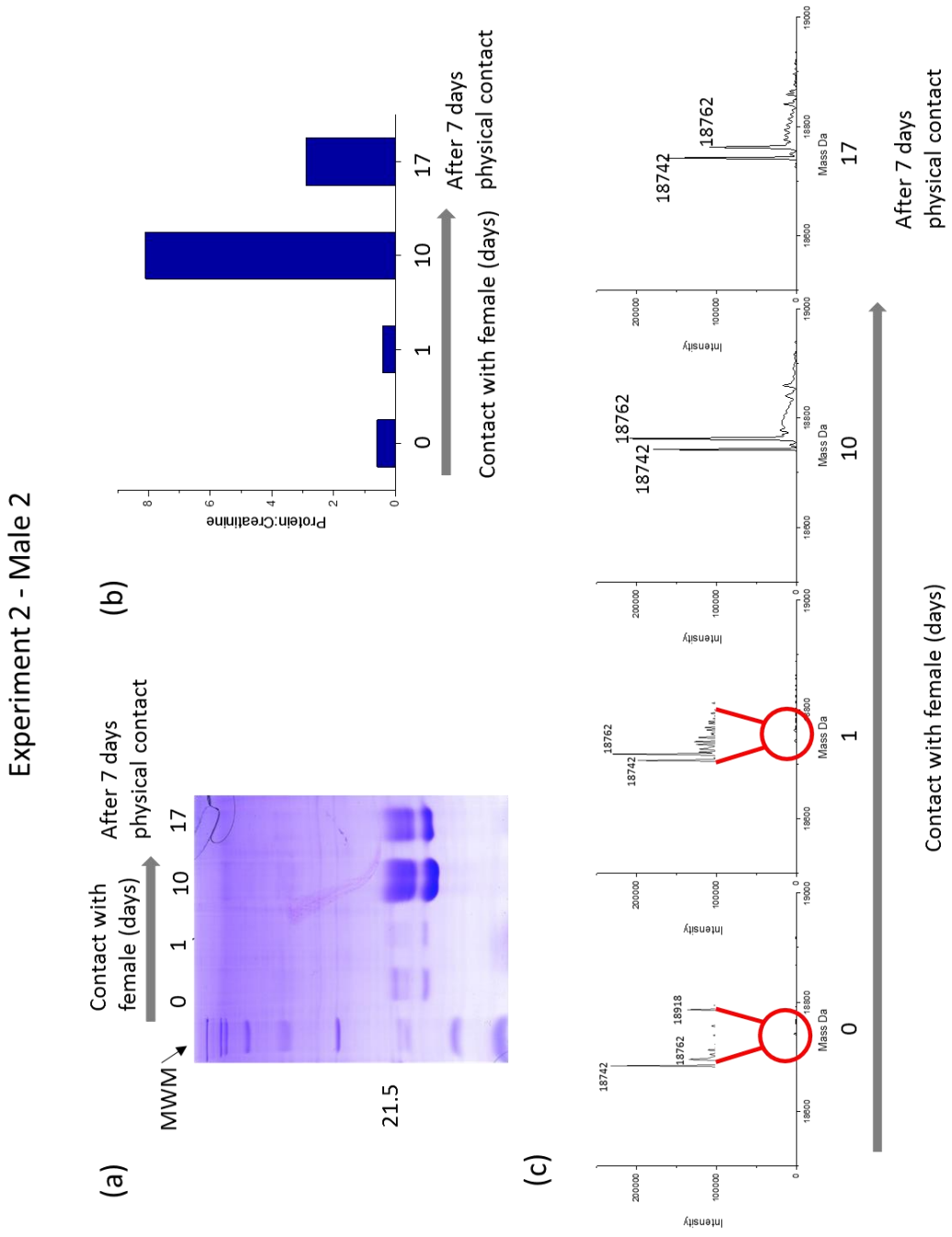

Experiment 2 - Male 2
(a)
MWM
Contact with female (days)
0
1
10
17
After 7 days physical contact
21.5
(b)
Protein:Creatinine
Contact with female (days)
After 7 days physical contact
(c)
Intensity
Mass Da
18742
18762
18918
Contact with female (days)
After 7 days physical contact

Figure 3.42 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and after being put together for breeding. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

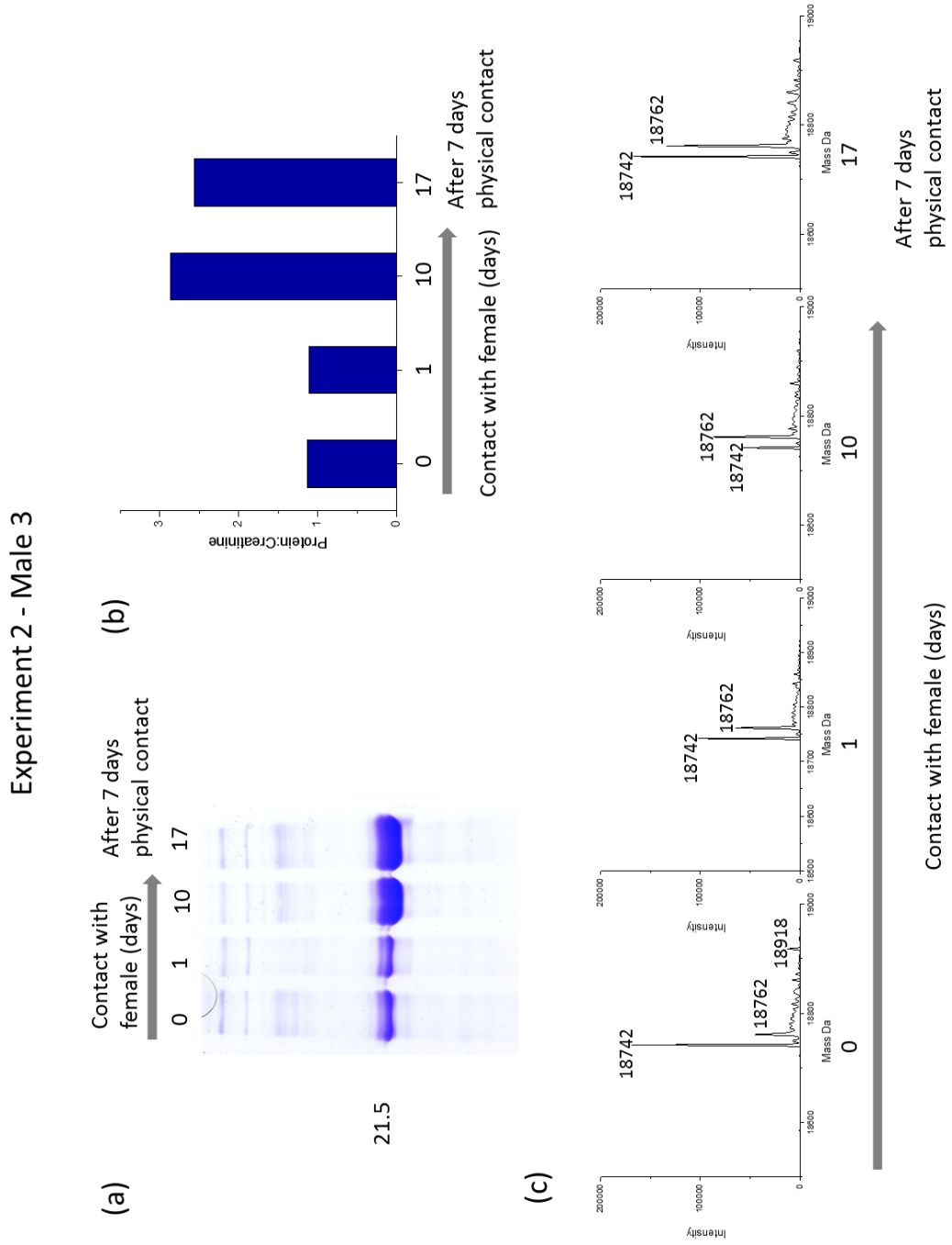

Experiment 2 - Male 3
(a)
Contact with female (days)
0
1
10
17
After 7 days physical contact
21.5
(b)
Protein:Creatinine
0
1
2
3
0
1
10
17
Contact with female (days)
After 7 days physical contact
(c)
Intensity
Mass Da
18742
18762
18918
0
1
10
17
Contact with female (days)
After 7 days physical contact

Figure 3.43 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and after being put together for breeding. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

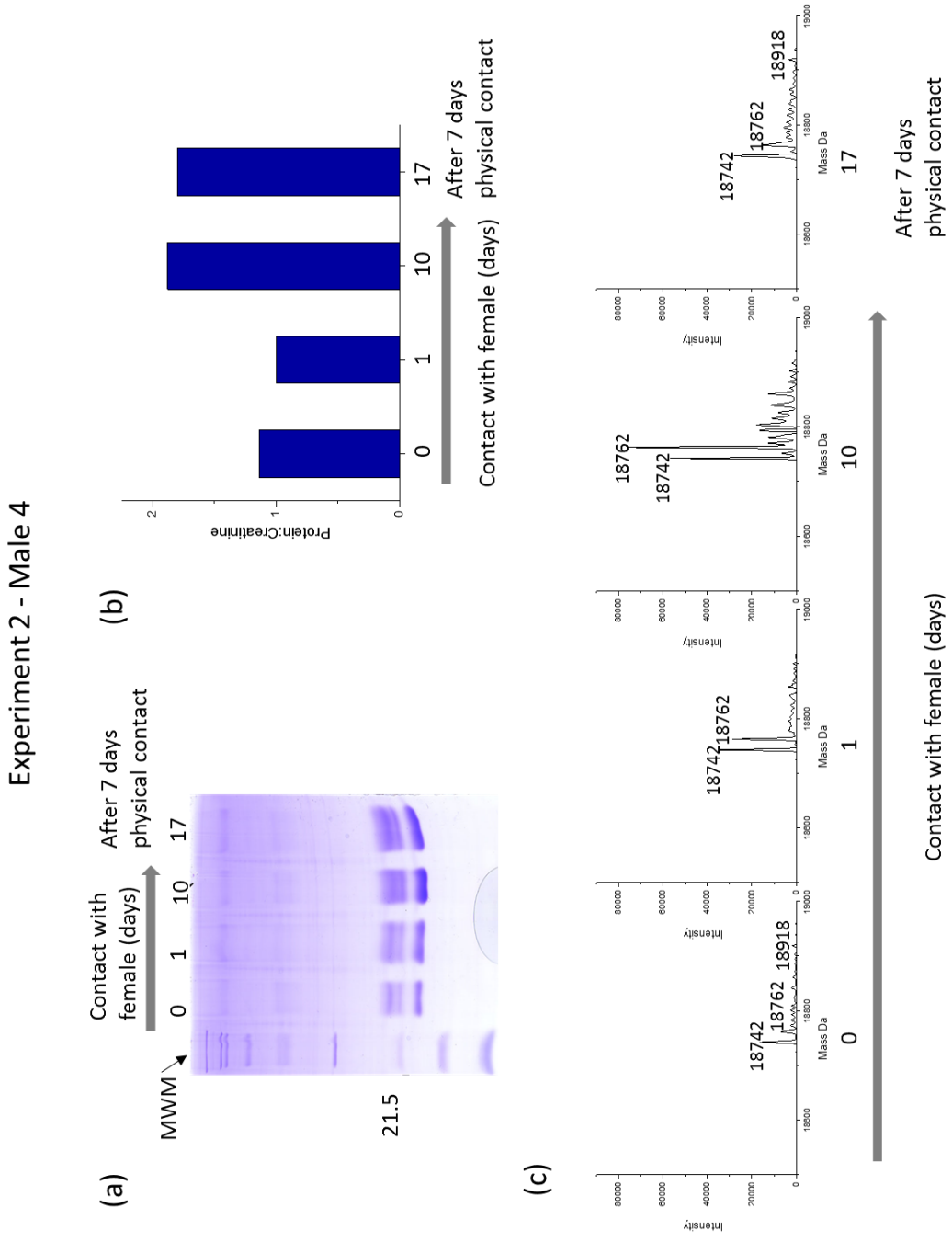

Experiment 2 - Male 4
(a)
MWM
Contact with female (days)
0
1
10
17
After 7 days physical contact
21.5
(b)
Protein:Creatinine
Contact with female (days)
After 7 days physical contact
(c)
Intensity
Mass Da
18742
18762
18918
Contact with female (days)
After 7 days physical contact

Figure 3.44 SDS-PAGE and ESI-MS analyses of male *M. spicilegus* urine.
Urine (5 µl) from a male *M. spicilegus,* taken prior to contact with the female, one day after contact began, after ten days of contact and after being put together for breeding. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution. (c) Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters).

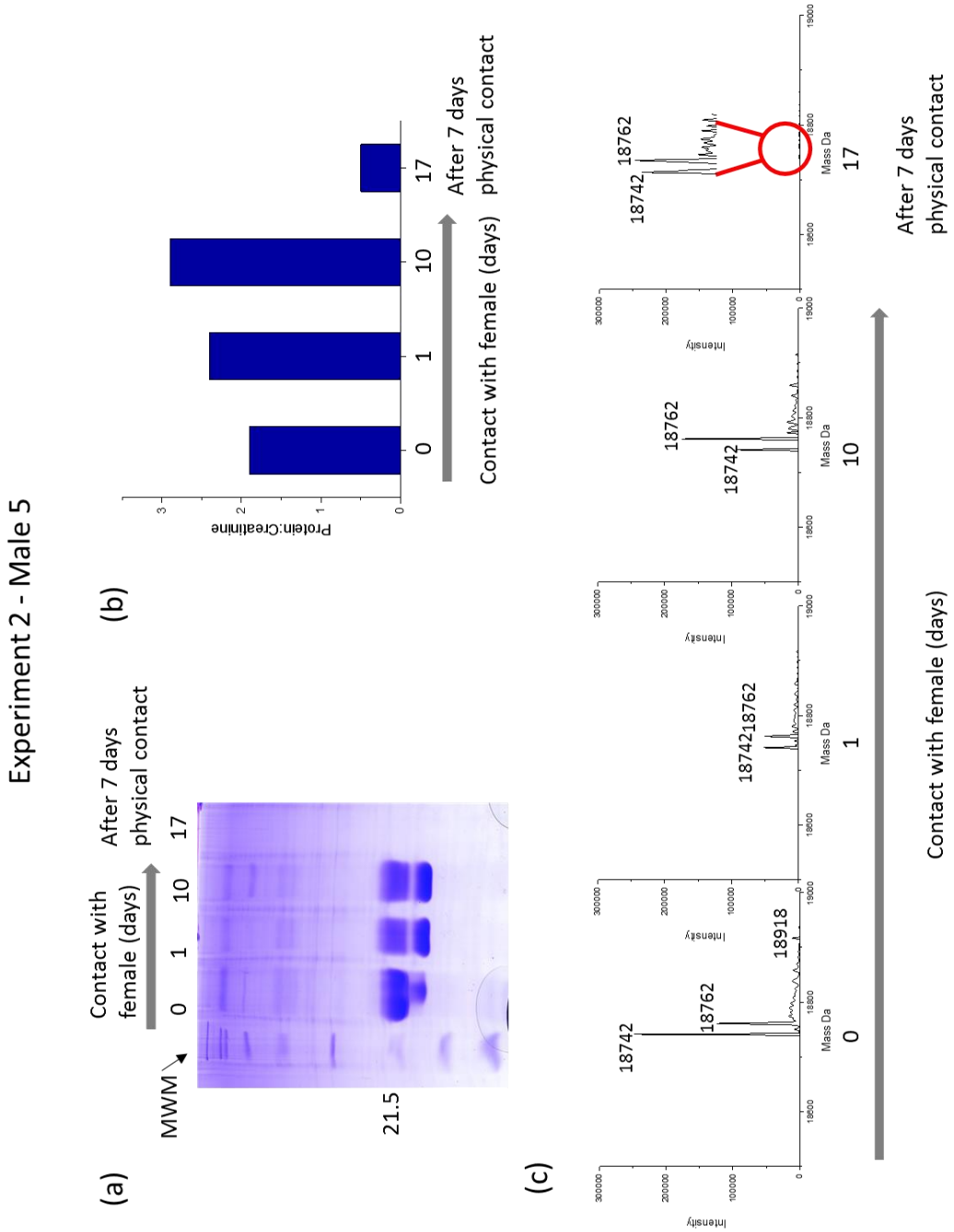
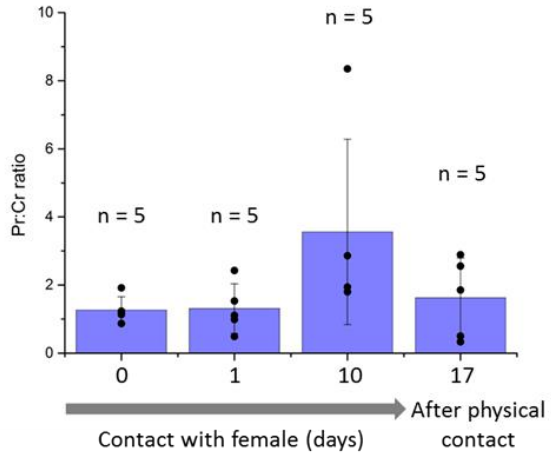
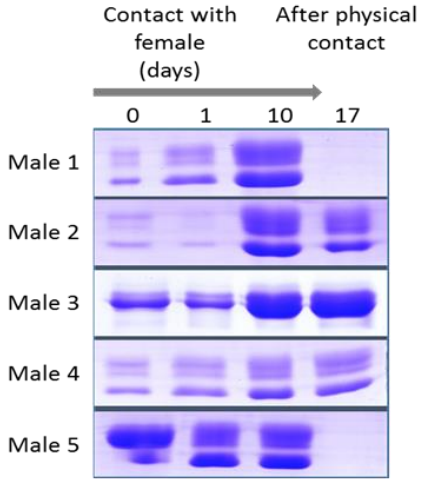
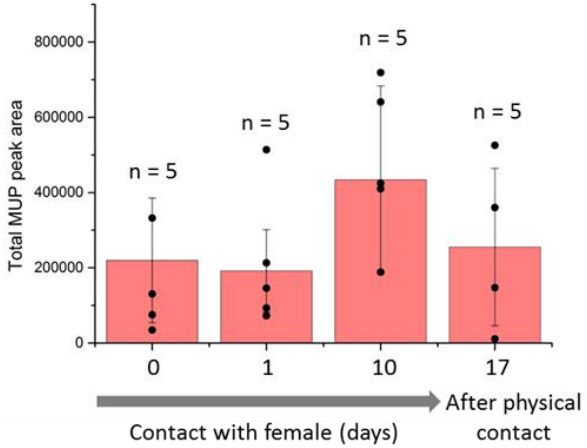
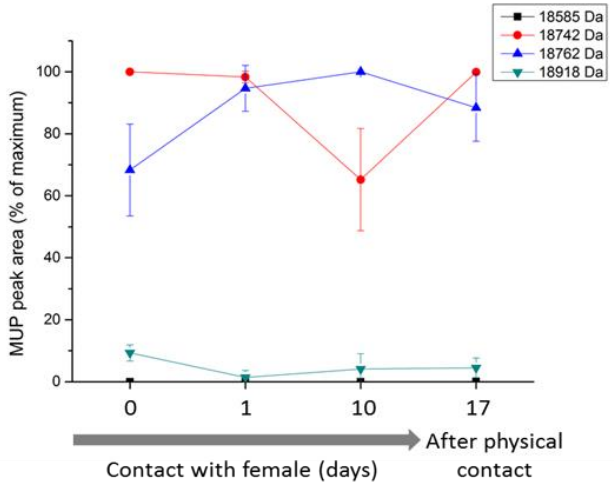

Experiment 2 - Male 5
(a)
MWM
Contact with female (days)
0
1
10
17
After 7 days physical contact
21.5
(b)
Protein:Creatinine
Contact with female (days)
After 7 days physical contact
(c)
Intensity
Mass Da
18742
18762
18918
Contact with female (days)
After 7 days physical contact

Figure 3.45 Summary of the analyses of *M. spicilegus* urine from male mice in Experiment 2.
Top left: The average protein:creatinine ratios calculated for the urine samples from each of the five males at each sample collection point (error bars ± SD). Data points for each sample are shown. Top right: The ~20 kDa region (MUP bands) of the SDS-PAGE analysis of the urine from each of the five males. Bottom left: The average total peak area of all *M. spicilegus* MUP peaks in ESI-MS analysis (error bars ± SD). MUP peak areas (providing a relative quantification) observed in each of the ESI-MS spectra were calculated using MaxENT software and summed for each sample. Data points for each sample are shown. Bottom right: The average peak areas of each *M. spicilegus* MUP peak in the ESI-MS analysis (error bars ± SD). The peak area of each MUP peak in each spectrum were calculated using MaxENT software. For each male urine sample, the MUP peak areas were expressed as a percentage of the maximum MUP peak area.


Pr:Cr ratio
n = 5
n = 5
n = 5
n = 5
0
1
10
17
Contact with female (days)
After physical contact


Contact with female (days)
After physical contact
0
1
10
17
Male 1
Male 2
Male 3
Male 4
Male 5


Total MUP peak area
n = 5
n = 5
n = 5
n = 5
0
1
10
17
Contact with female (days)
After physical contact


18585 Da
18742 Da
18762 Da
18918 Da
MUP peak area (% of maximum)
0
1
10
17
Contact with female (days)
After physical contact

Experiment 3 – Contact and interactions between male/male pairs

In the final experiment, six non-sibling male/male pairs were placed in large split cages, where a mesh barrier was placed between them which they could contact each other through, but not gain full access to each other (Figure 3.46). The male/male pairs were housed in this setting for 12 days, however only three pairs were kept in this setting successfully for this time – three of the pairs managed to displace the barriers between them at various times. For the remaining three pairs, after 12 days of being housed in contact, they were observed interacting with each other with no barrier between them, to make note of any dominant behaviour exhibited by any of the animals. These controlled interactions took place for ten minutes each day for the following three days. Following the final interaction, the bedding from their cages was removed and replaced with paper to collect and view their scent marks, with the contact barrier remaining between them. The urine of these three male/male pairs was sampled using the recovery method on the day prior to them coming into contact, one day after they began contact, 12 days after they began contact (prior to controlled interactions) and on the final day (after the controlled interactions). The aim of this experiment was to determine whether contact with an unknown male would cause a male to alter his own MUP expression, possibly to express dominance, and the behaviour observed and noted during the interactions would allow the determination of any link between behaviour and MUP expression. The collection of scent marks would allow us to see any further link between any dominant behaviour, dominant scent marking and MUP expression patterns, assuming the scent marking behaviour of *M. spicilegus* was the same as *M. m. domesticus* (Rich and Hurst 1998; 1999).

All urine samples were frozen upon collection until further analysis. All samples recovered from male mice were analysed for the presence and relative abundance of MUPs using SDS-PAGE analysis, and samples were then analysed using ESI-MS to provide an insight into the relevant amounts of each MUP expressed at each stage of the experiment. Protein and creatinine assays were carried out to confirm the changing concentration of protein in the urine. Scent marking paper was immediately collected, viewed under UV light and photographed. The behaviour observed from all mice during the interactions was noted throughout each interaction.

In the first male/male pair, SDS-PAGE gels indicated that animal A expressed a similar amount of MUP during the contact period as it did prior to contact to the

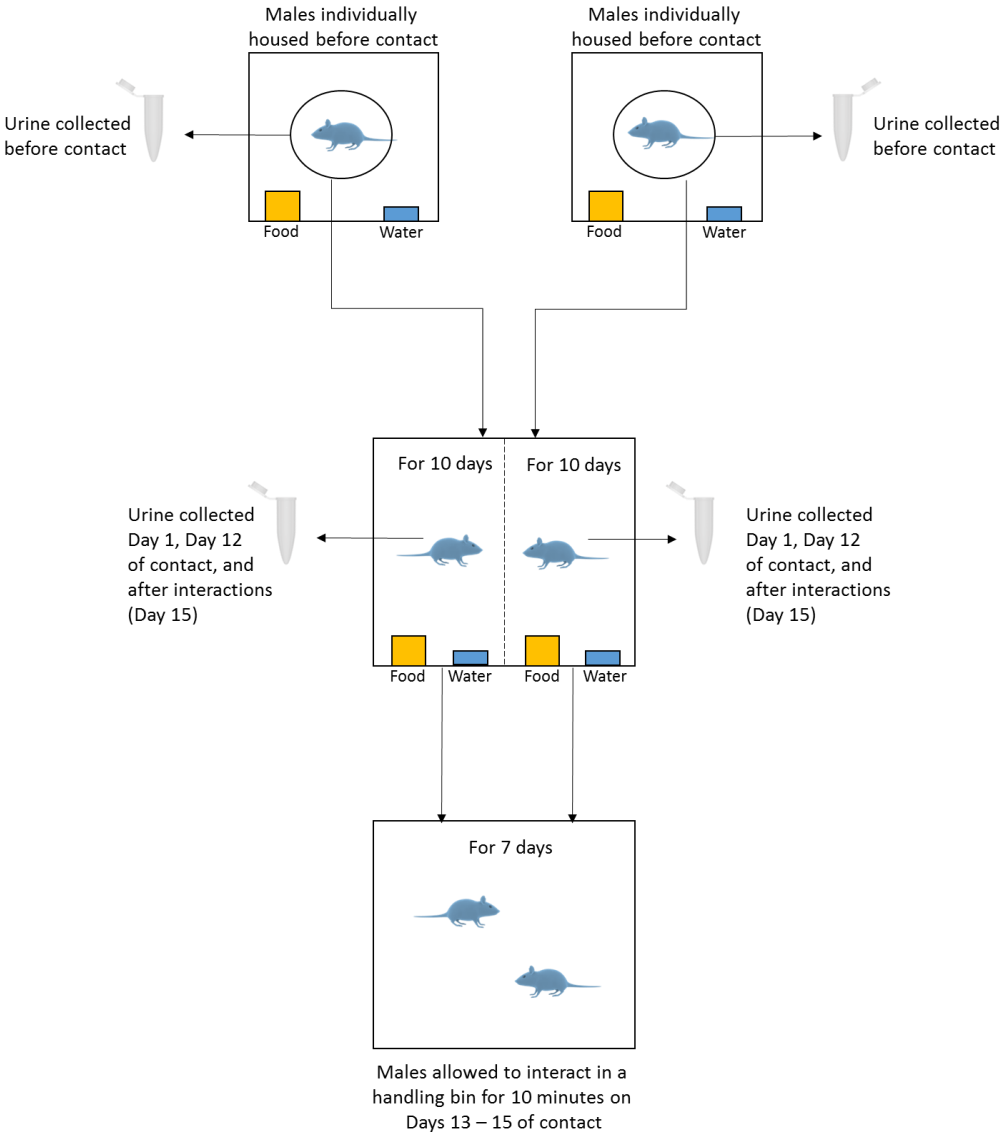


Figure 3.46 Set-up of Experiment 3.
Set-up of Experiment 3 – contact and interaction between male/male pairs. Urine samples were taken from two non-sibling singly-housed males before the pair were housed in a split cage, separated by a mesh barrier which allowed contact between the males, but not full physical access to one another. Pairs were housed in the split cages for 12 days, with urine collected from mice on days 1 and 12 of contact. On days 13 – 15 of contact, males were allowed to interact in a handling bin for 10 minutes daily. A final urine sample and scent marks were taken from each animal on the final day of the experiment (Day 15).

unknown male. Animal B expressed less MUP than animal A in general, but expressed an increasing amount of MUP after 12 days of contact with animal A. This was confirmed by the calculated protein:creatinine ratios in each of these samples (Figure 3.47). ESI-MS showed that the two males had similar MUP expression profiles prior to their contact, and whilst the MUP expression pattern of animal A remained fairly similar up until the interactions took place, animal B appeared to express increased amounts of the 18762 Da MUP after 12 days of contact (Figure 3.48). During the first interaction, both A and B exhibited investigative behaviour (sniffing, approaching) to a similar extent, but animal B exhibited aggressive behaviour (chasing, attempting to fight) towards animal A towards the end of the ten minute period. During the second and third interactions, both A and B exhibited significant aggressive behaviour towards each other and had to be split up on both occasions (Supplementary Material). When returned to their barrier-split cage, both mice deposited many scent marks, with animal A marking close to the barrier and animal B marking away from the barrier (Figure 3.49). The SDS-PAGE analysis showed that after interactions, animal A (who previously expressed less MUP) expressed significantly more MUP, and animal B, the converse. This was confirmed by ESI-MS, with total MUP peak area increasing significantly in Animal A after interactions, and decreasing in Animal B (Figure 3.50). Animal A expressed more of the 18742 Da and 18762 Da MUPs relatively after interactions, and animal B expressed less of these MUPs (Figure 3.50).

In the second male/male pair, SDS-PAGE gels indicated that both animals C and D expressed similar amounts of MUP during the contact period as they did prior to contact with the unknown male (Figure 3.51). Animal C expressed slightly less MUP after 12 days of contact, but MUP expression increased after the controlled interactions. The levels of expression were confirmed by the calculated protein:creatinine ratios in each of these samples (Figure 3.51). MUP expression patterns for both animals were very similar to each other throughout the experiment, with little change in MUP expression in each urine sample taken – at no point was the expression of the 18762 Da MUP increased significantly, unlike the first male/male pair (Figures 3.52 and 3.54). During the first interaction, both animals C and D exhibited investigative behaviour initially, but animal C approached and chased animal D, with animal D retreating and trying to avoid animal C. No aggressive behaviour was observed at any point. During the second interaction, investigative behaviour was observed in both animals, with animal C showing more than animal D. Animal C was seen to try and climb on top of animal D, and animal D kept avoiding

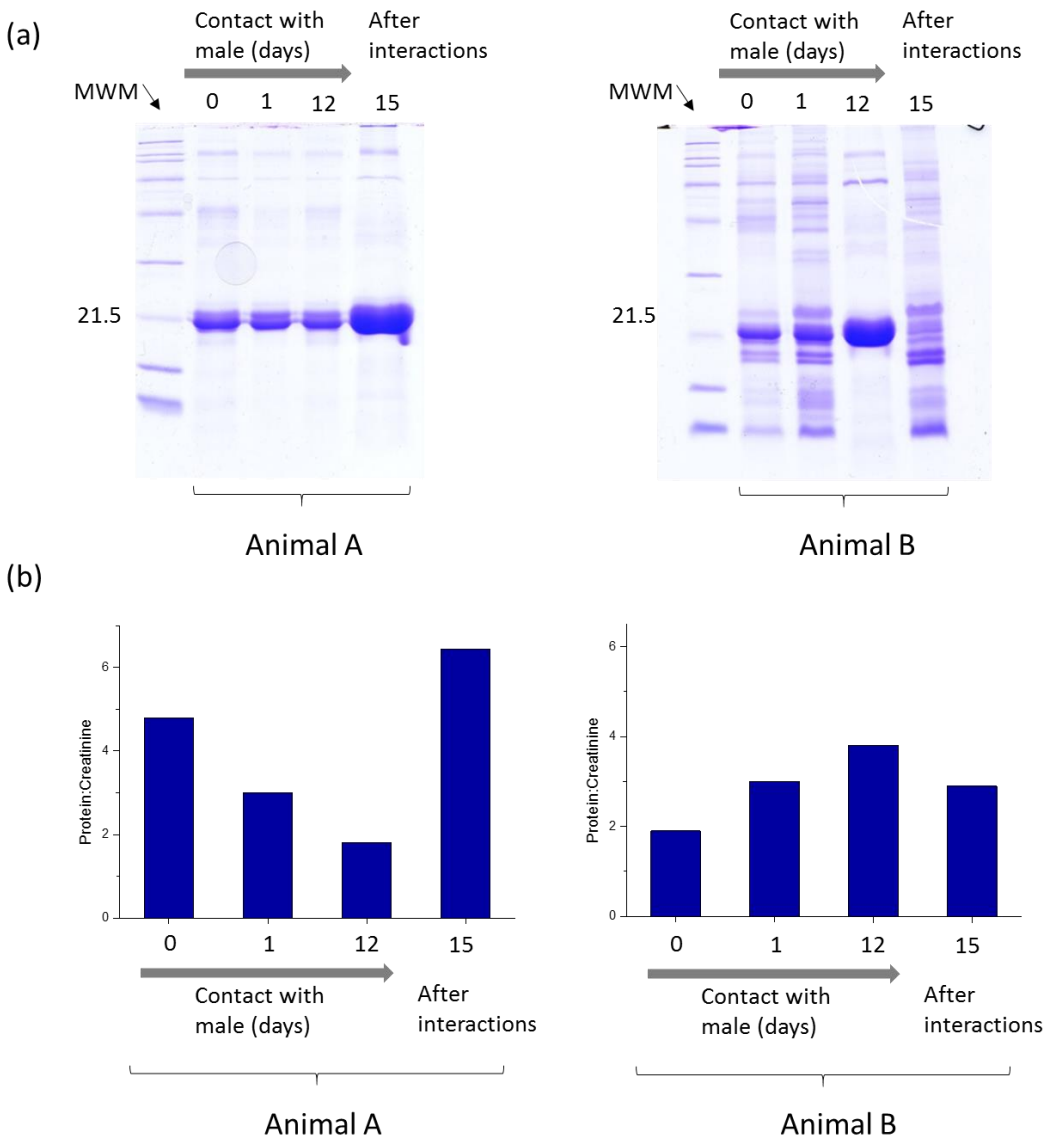


Figure 3.47 SDS-PAGE analysis and protein:creatinine ratios of male *M. spicilegus* urine from the first male/male pair.
Urine (5 µl) from the first male/male *M. spicilegus* pair, (male A and male B), taken prior to contact with the unfamiliar male, one day after contact began, after 12 days of contact and after being put together for controlled interactions. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution.

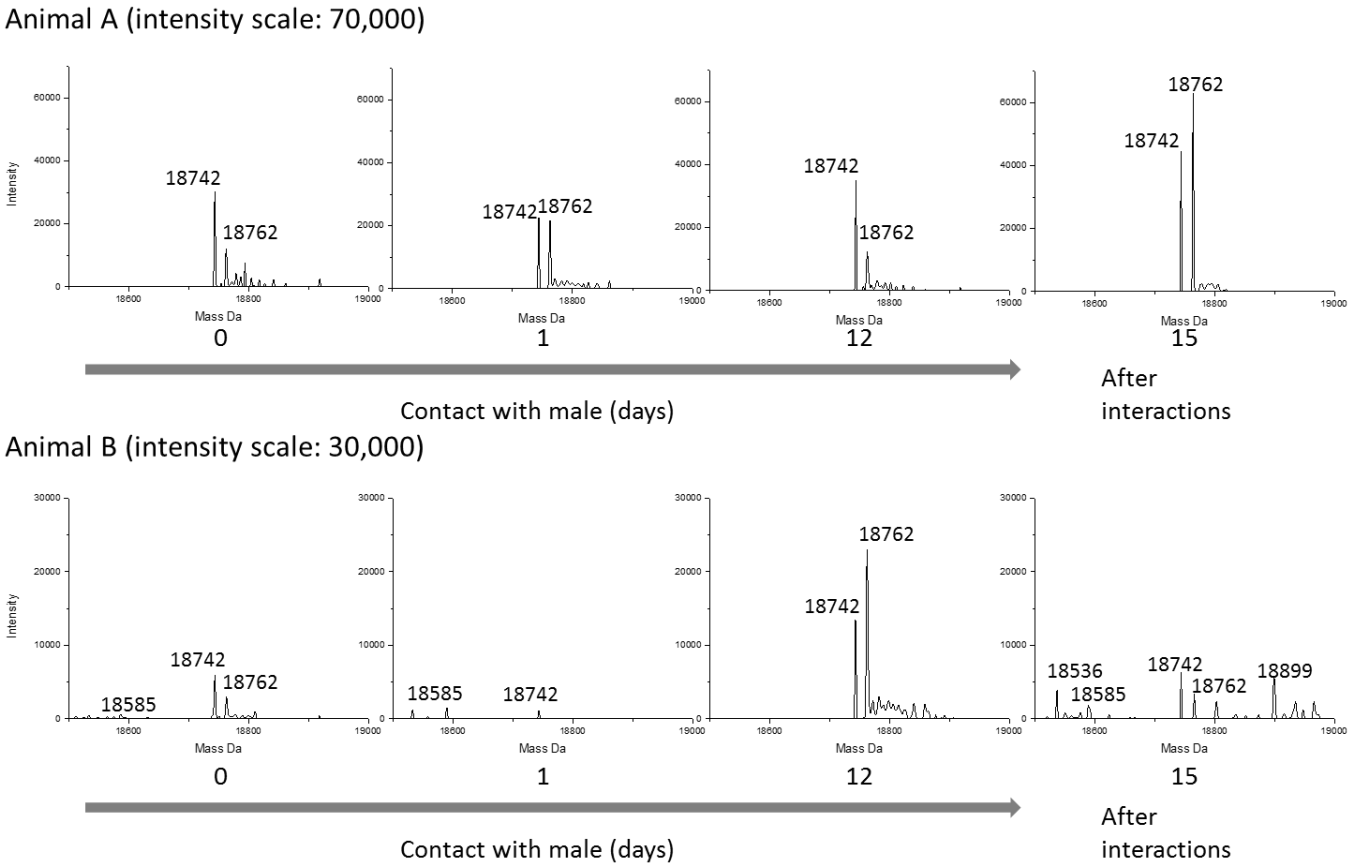


Figure 3.48 ESI-MS analysis of male *M. spicilegus* urine from the first male/male pair.
Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). All spectra obtained for each male is on the same scale to visualise differences in MUP expression at each stage of the experiment. Differences in scale between animals A and B are noted above the spectra.

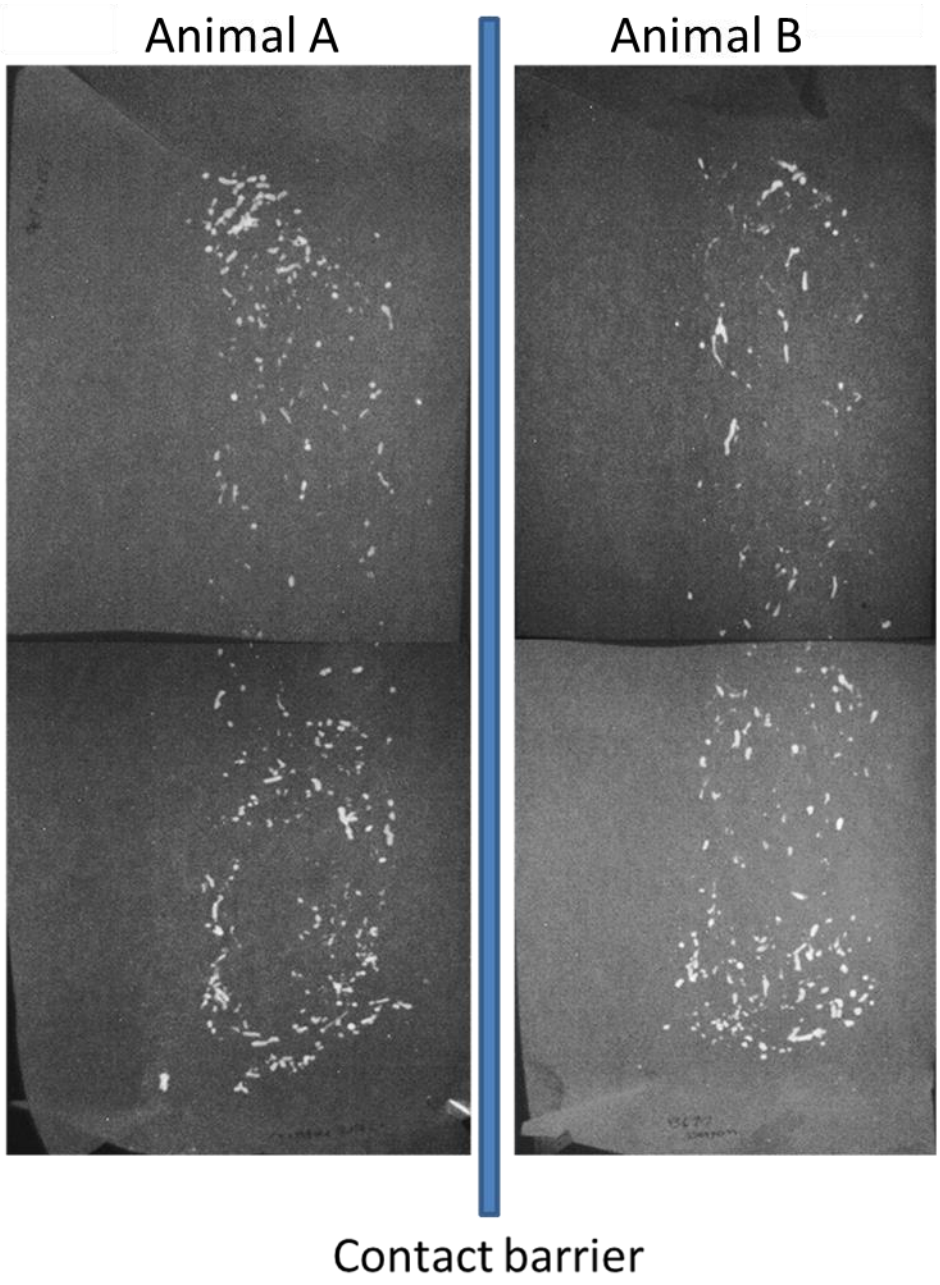


Figure 3.49 Photographs of male *M. spicilegus* scent marks deposited by the first male/male pair.
Males A and B were placed on scent marking paper in their barrier-split cage for ten minutes. The paper was recovered and scent marks were visualised under UV light. Deposited scent marks fluoresce under UV light and are the light grey areas seen on the paper.

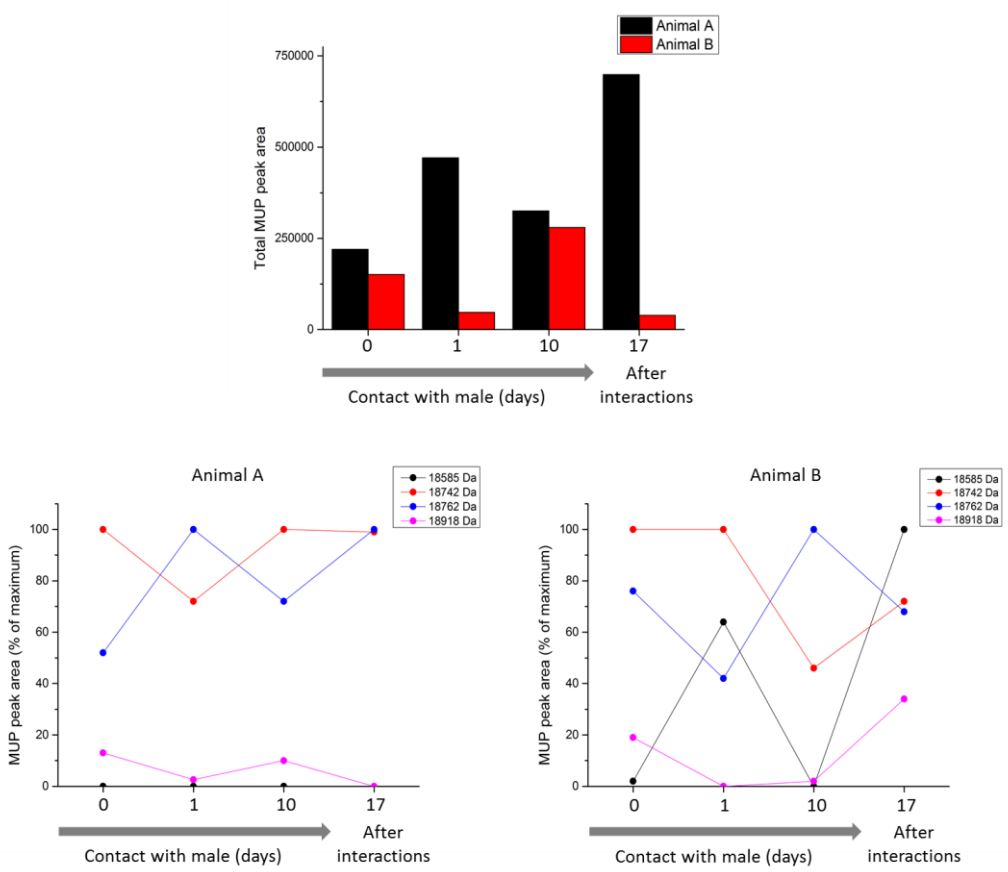


Figure 3.50 Summary of the analyses of *M. spicilegus* urine from the first male pair in Experiment 3.
Top: MUP peak areas (providing a relative quantification) observed in each of the ESI-MS spectra were calculated using MaxENT software and summed for each sample. Bottom: The average peak areas of each *M. spicilegus* MUP peak in the ESI-MS analysis of urine from Animal A and Animal B. The peak area of each MUP peak in each spectrum were calculated using MaxENT software. For each urine sample, the MUP peak areas were expressed as a percentage of the maximum MUP peak area.

animal C. During the final interaction, both mice showed a little investigative behaviour but largely ignored each other. When returned to their barrier-split cage, neither animal scent marked (Figure 3.53). The lack of significant dominant behaviour during interactions, lack of scent marking and little change in MUP output and expression indicates that changes in MUP output may be related to dominant, aggressive behaviour.

In the third male/male pair, SDS-PAGE gels indicated that animal E expressed very little MUP prior to contact with an unknown male, and increased MUP expression throughout contact. MUP expression was increased further after interactions with animal F. Animal F expressed some MUP prior to contact, and showed little difference in MUP expression throughout contact and after interactions. Animal F had notably higher levels of albumin in all urine samples (Figure 3.55). ESI-MS spectra show that MUP expression patterns by animal E remain fairly similar throughout contact, with slight variation in the relative intensity of the 18762 Da MUP (Figures 3.56 and 3.58). MUP expression patterns by animal F also remain fairly similar throughout contact, with the 18762 Da MUP more intense than the 18742 Da MUP in each urine sample (Figure 3.56). Relative quantification of the MUPs in samples from animal F shows that the expression of the 18742 Da and 189818 Da MUPs decreases upon contact with animal E (Figure 3.58). During all three interactions, neither animal displayed any aggressive or avoidance behaviour, with investigative behaviour displayed during the first interaction, and remaining in close proximity to each other for the remainder of the interactions. When returned to their barrier-split cage, animal E scent marked fairly significantly, whilst animal F did not (Figure 3.57). This, along with animal E expressing more urinary MUP after the controlled interactions (Figure 3.58), suggests a link between MUP output and scent marking behaviour. The fact that MUP expression patterns in both animals did not change significantly throughout the contact period/after interactions, and that harmonious behaviour during the interactions was observed, suggests that changes in the total amount of MUP expressed may be related to dominant behaviour.

There have not been enough samples to determine a definitive link between MUP expression patterns and behavioural characteristics observed between unfamiliar male/male pairs, but the fact that the only pair out of the three to show any significant changes in the total amount of MUP expressed were the only pair to display aggressive behaviour, suggests that the up-regulation of MUPs may be linked to the assertion of dominance, whilst the down-regulation of these MUPs may be linked to

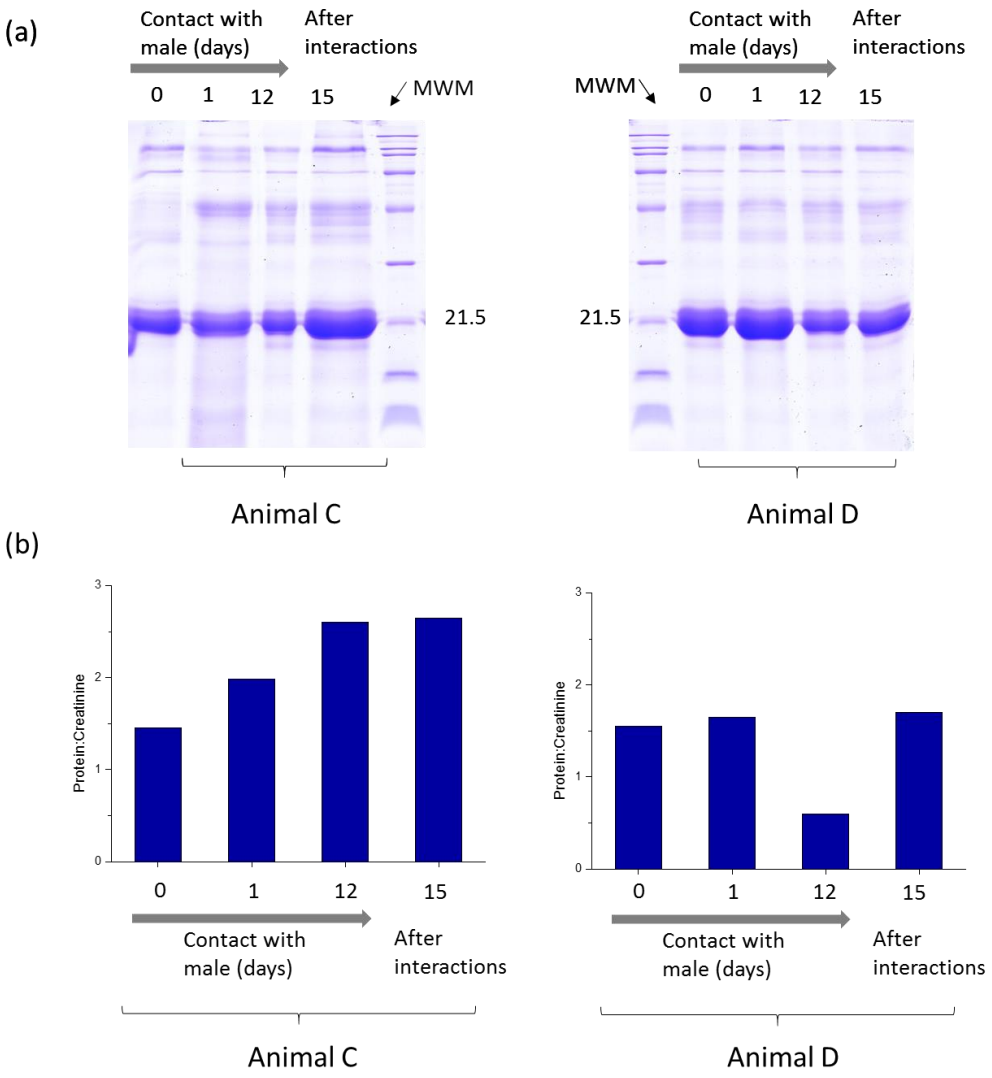


Figure 3.51 SDS-PAGE analysis and protein:creatinine ratios of male *M. spicilegus* urine from the second male/male pair.
Urine (5 µl) from the second male/male *M. spicilegus* pair, (male C and male D), taken prior to contact with the unfamiliar male, one day after contact began, after 12 days of contact and after being put together for controlled interactions. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution.

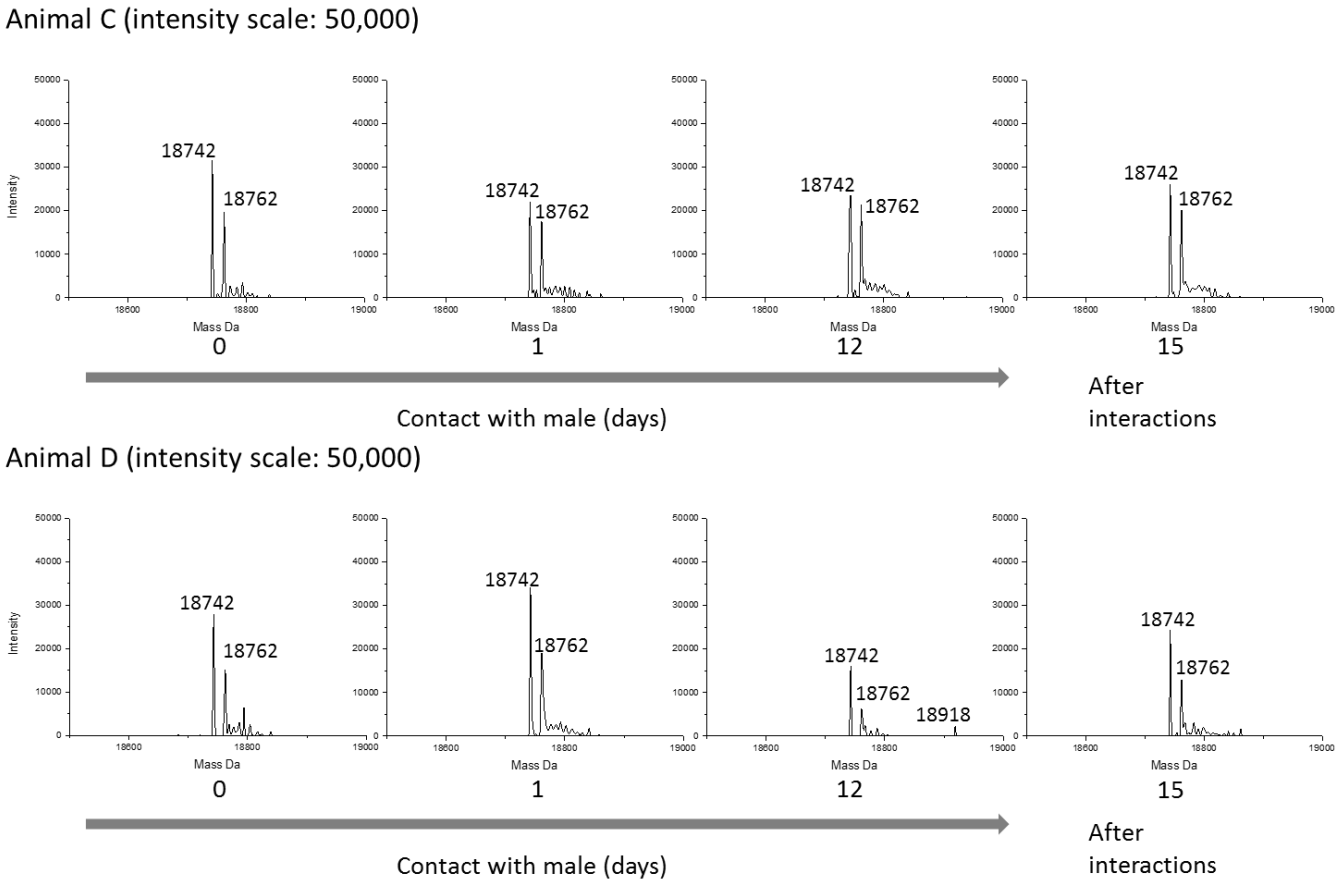


Figure 3.52 ESI-MS analysis of male *M. spicilegus* urine from the second male/male pair.
Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). All spectra obtained for each male is on the same scale to visualise differences in MUP expression at each stage of the experiment. Differences in scale between animals A and B are noted above the spectra.

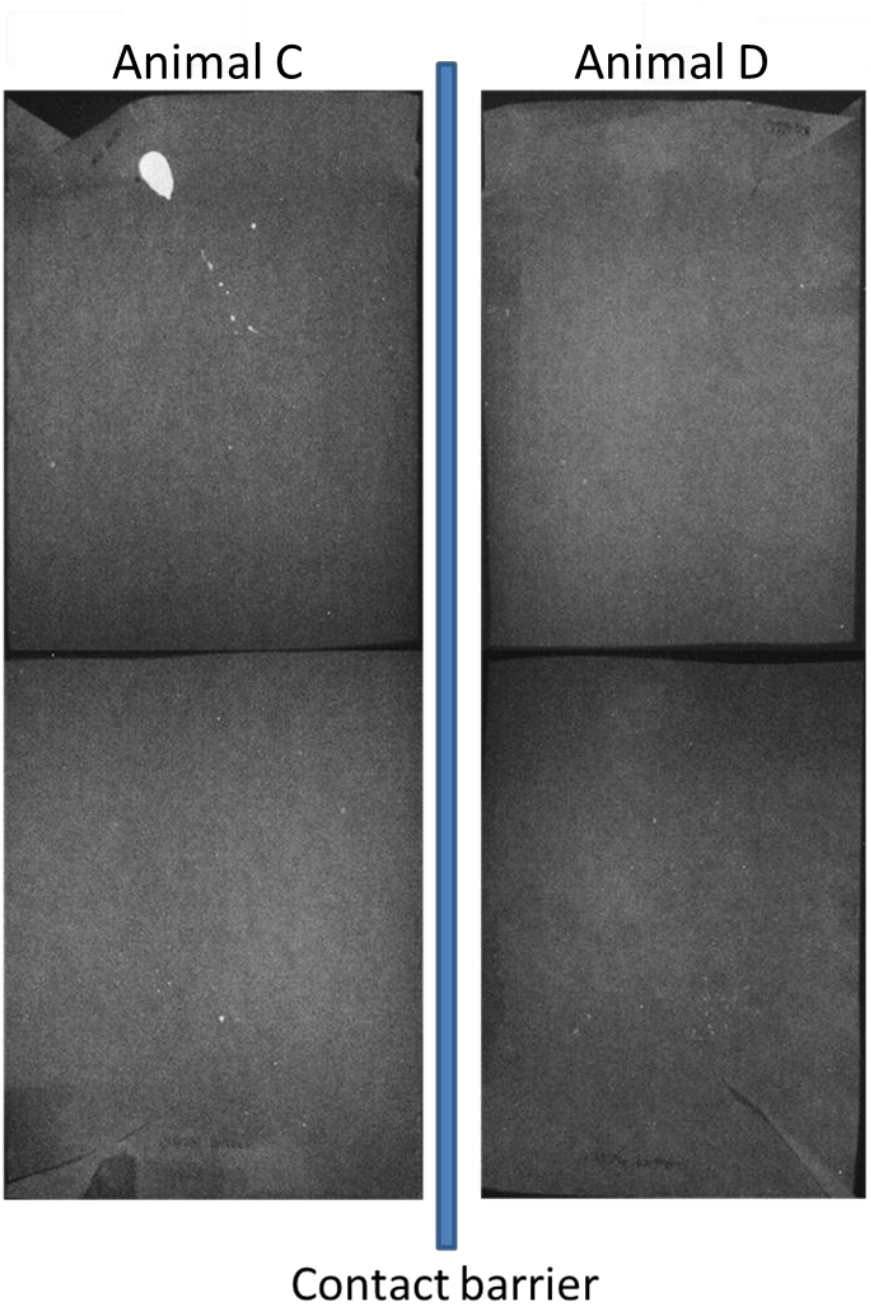


Figure 3.53 Photographs of male *M. spicilegus* scent marks deposited by the second male/male pair.
Males C and D were placed on scent marking paper in their barrier-split cage for ten minutes. The paper was recovered and scent marks were visualised under UV light. Deposited scent marks fluoresce under UV light and are the light grey areas seen on the paper.

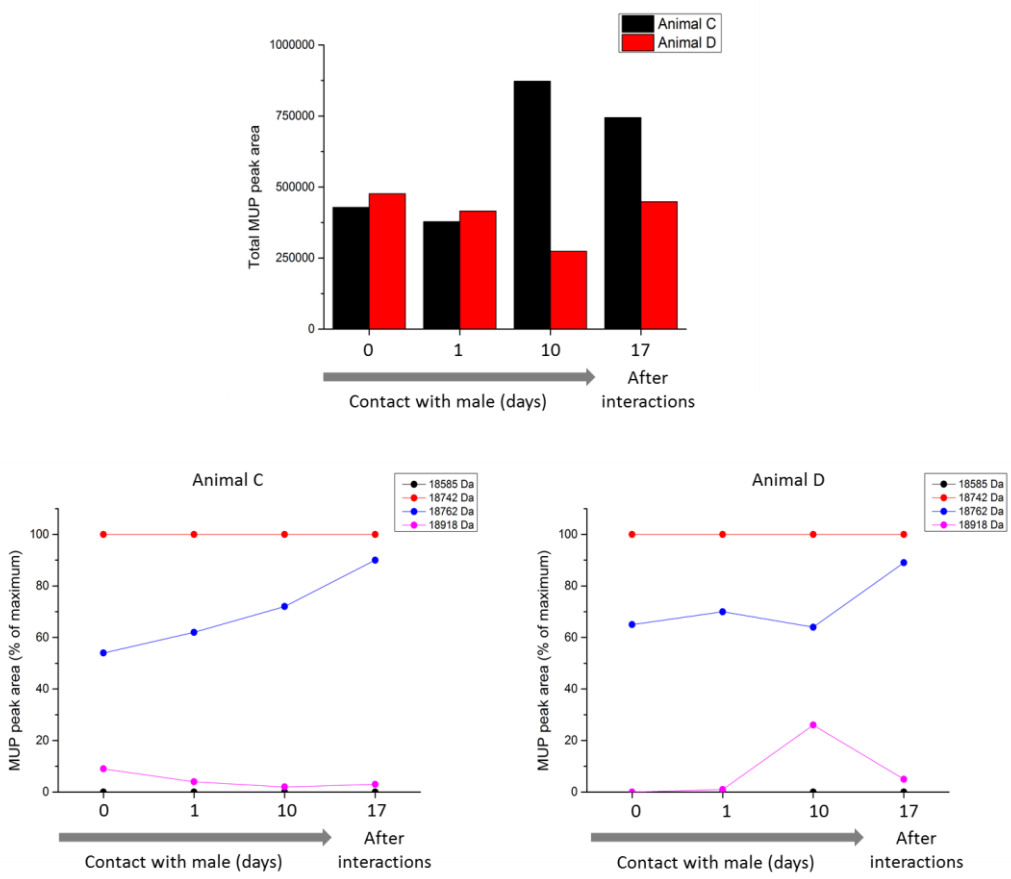


Figure 3.54 Summary of the analyses of *M. spicilegus* urine from the second male pair in Experiment 3.
Top: MUP peak areas (providing a relative quantification) observed in each of the ESI-MS spectra were calculated using MaxENT software and summed for each sample. Bottom: The average peak areas of each *M. spicilegus* MUP peak in the ESI-MS analysis of urine from Animal A and Animal B. The peak area of each MUP peak in each spectrum were calculated using MaxENT software. For each urine sample, the MUP peak areas were expressed as a percentage of the maximum MUP peak area.

submissiveness. Male MUP expression, particularly the 18762 Da MUP, was also seen to increase with contact with a female. This could suggest a link between MUP expression (particularly the 18762 Da MUP) and the attempts of the male to increase his attractiveness to a female – possibly through the indication of dominance.

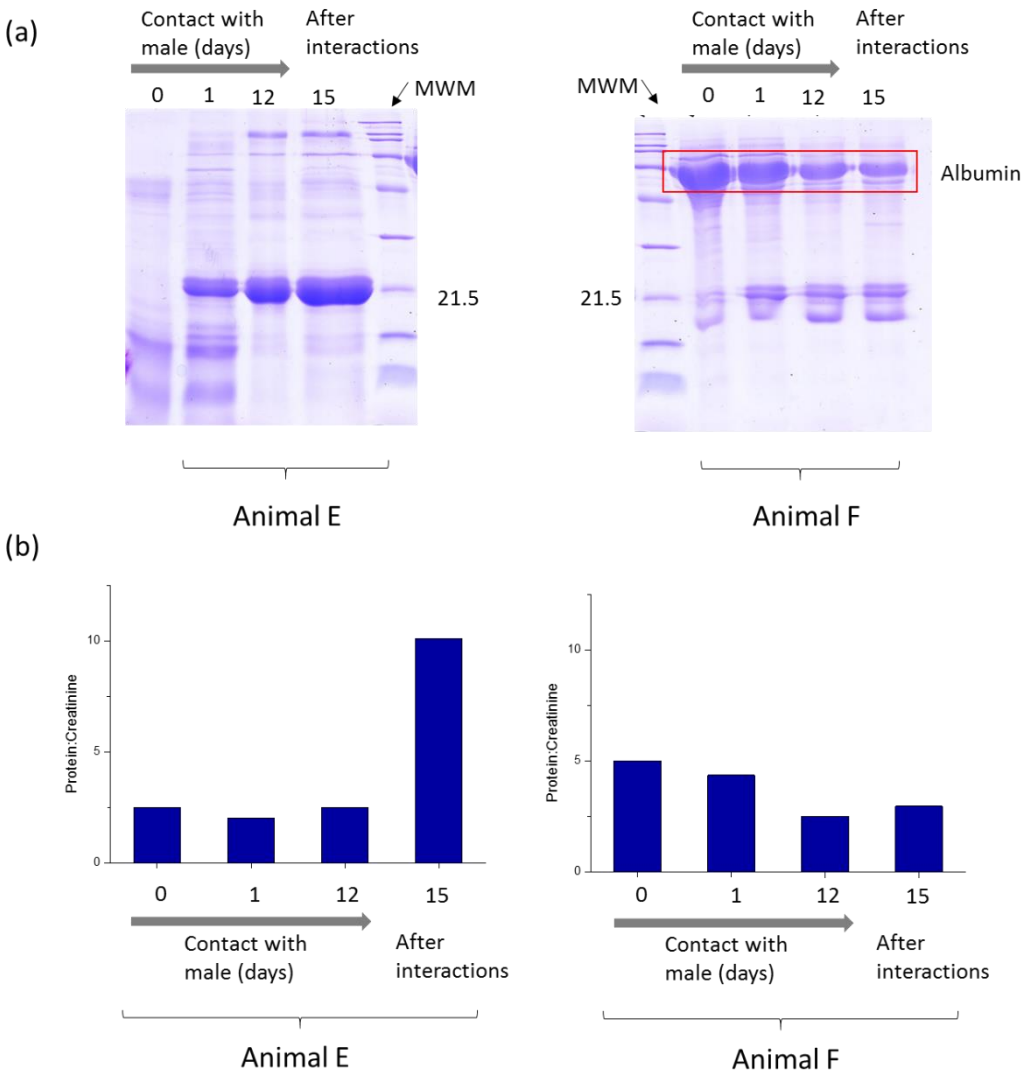


Figure 3.55 SDS-PAGE analysis and protein:creatinine ratios of male *M. spicilegus* urine from the third male/male pair.
Urine (5 µl) from the third male/male *M. spicilegus* pair, (male E and male F), taken prior to contact with the unfamiliar male, one day after contact began, after 12 days of contact and after being put together for controlled interactions. (a) Urine was mixed 1:1 with sample buffer, loaded onto a 15% SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. MUPs are seen at approximately 20 kDa. (b) Urinary protein and creatinine concentrations were measured as described in Chapter 2, and protein:creatinine ratio calculations to correct for urine dilution.

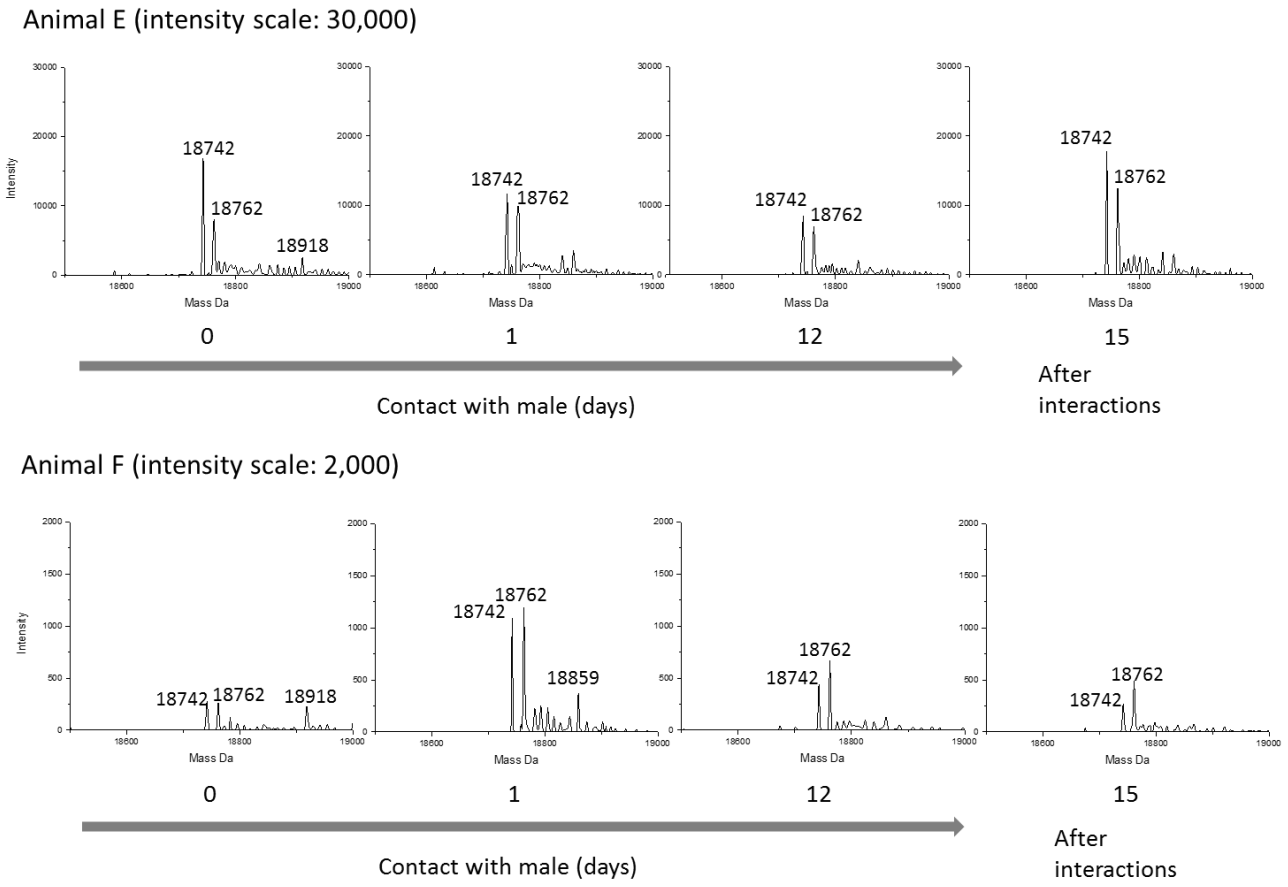


Figure 3.56 ESI-MS analysis of male *M. spicilegus* urine from the third male/male pair.
Accurate molecular weights of the MUPs present in the SDS-PAGE bands were identified by ESI-MS. Urine samples were diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). All spectra obtained for each male are on the same scale to visualise differences in MUP expression at each stage of the experiment. Differences in scale between animals A and B are noted above the spectra.

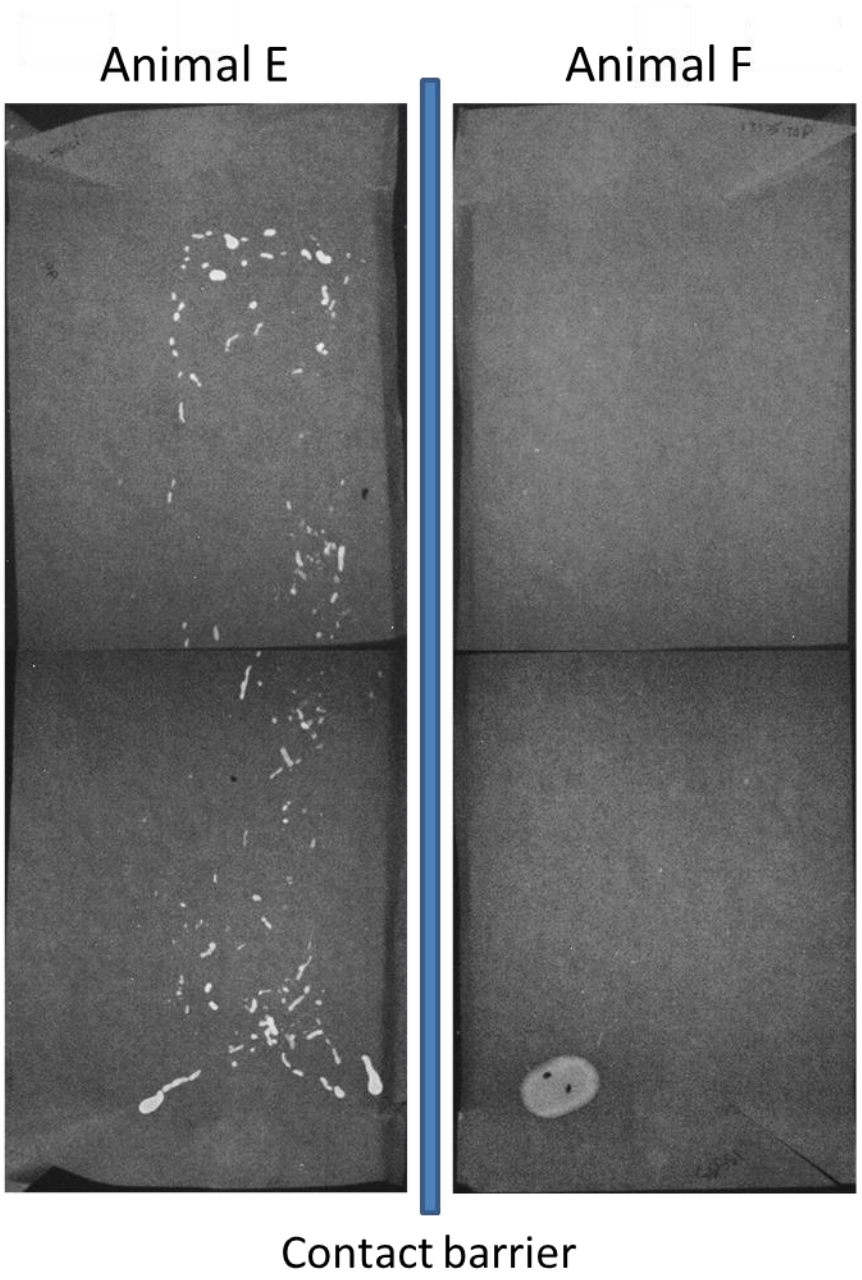


Figure 3.57 Photographs of male *M. spicilegus* scent marks deposited by the third male/male pair.
Males E and F were placed on scent marking paper in their barrier-split cage for ten minutes. The paper was recovered and scent marks were visualised under UV light. Deposited scent marks fluoresce under UV light and are the light grey areas seen on the paper.

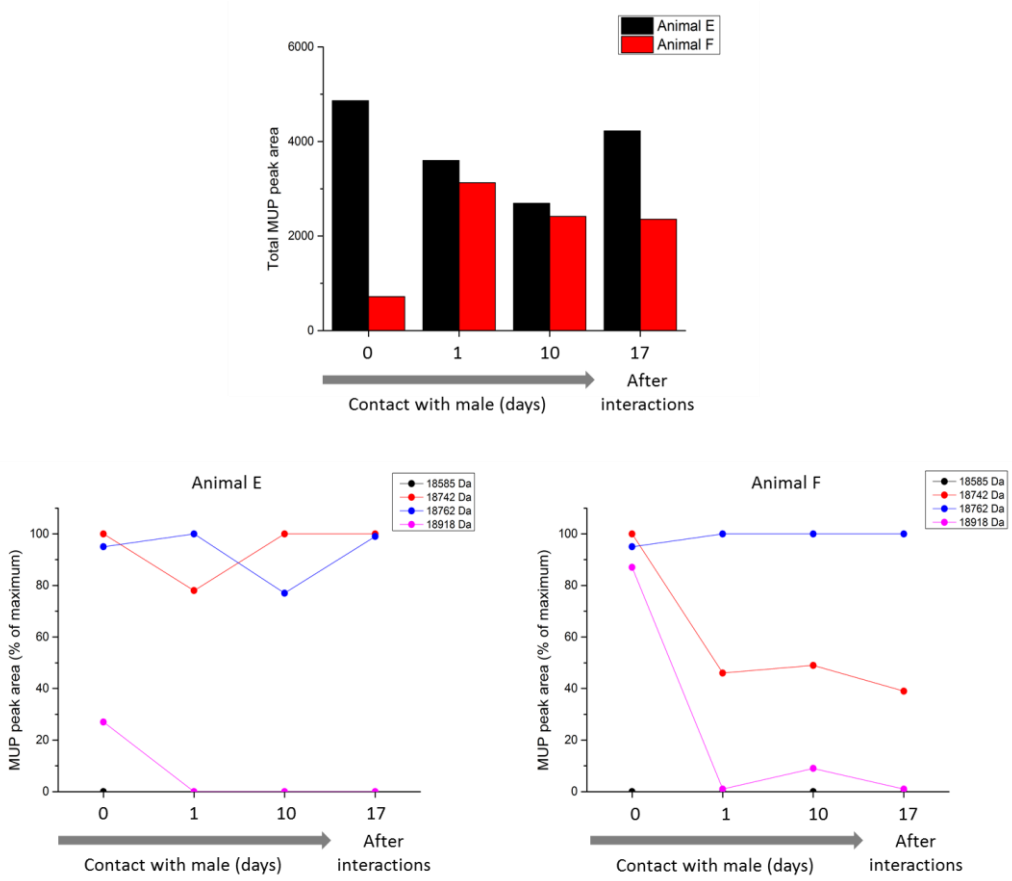


Figure 3.58 Summary of the analyses of *M. spicilegus* urine from the third male pair in Experiment 3.
Top: MUP peak areas (providing a relative quantification) observed in each of the ESI-MS spectra were calculated using MaxENT software and summed for each sample. Bottom: The average peak areas of each *M. spicilegus* MUP peak in the ESI-MS analysis of urine from Animal A and Animal B. The peak area of each MUP peak in each spectrum were calculated using MaxENT software. For each urine sample, the MUP peak areas were expressed as a percentage of the maximum MUP peak area.

3.3 Conclusions

Four major urinary proteins were identified and sequenced in the urine of *M. spicilegus* mice in the absence of genomic data. Male mice expressed more MUP in their urine than female mice, with the amount of MUP expressed by different males varying quite significantly – urinary protein:creatinine mass concentration ratios in singly housed males with no previous adult contact with other males or females ranged from 0.3 – 1.3; the urinary protein:creatinine concentration ratios in males contacting/interacting with other males or females ranged from 0.4 – 12.5. The amount of MUP expressed appeared almost identical amongst females.

In male mice, MUP expression patterns also varied between individuals. All male mice expressed the 18742 Da and 18762 Da MUPs in their urine; some also expressed the 18585 Da MUP, and some expressed the 18918 Da MUP (Figure 3.59). All female urine samples that contained enough MUP for ESI-MS analysis only contained the 18918 Da MUP. There was no evidence of the 18742 Da, 18762 Da and 18585 Da MUPs in any female urine sample (Figure 3.5). For this reason, it was likely that these three MUPs were male specific, whilst the 18918 Da, seen in both male and female urine samples, was not sex specific. The molecular weights of the four identified *M. spicilegus* MUPs did not match the molecular weights of any previously identified *M. m. domesticus* MUPs, therefore further analysis was needed to identify the sequences of the *M. spicilegus* MUPs.

Further mass spectrometric analysis confirmed numerous similarities between the peptide mass fingerprints of *M. spicilegus* MUPs and known *M. m. domesticus* MUPs (Figures 3.6 – 3.9). In the peptide mass fingerprints of the two distinct MUP bands observed in the SDS-PAGE gels of male urine, there were many similarities between the two bands, and upon database searching, many of the observed peptide masses matched to those of central *M. m. domesticus* MUPs. In the peptide mass fingerprints of the MUP band observed in the SDS-PAGE gels of female urine, there were minimal similarities between the female and male spectra. Database searches matched many of the peptides in the female MUP band to those of peripheral *M. m domesticus* MUPs. From these analyses, it was determined that the non sex specific 18918 Da MUP was similar in sequence to peripheral *M. m. domesticus* MUPs, and that the male specific 18742 Da, 18762 Da and 18585 Da MUPs were similar in sequence to central *M. m. domesticus* MUPs.

The male specific MUPs 18742 Da and 18762 Da were successfully separated using strong anion exchange chromatography for *de novo* sequencing analysis. Both of

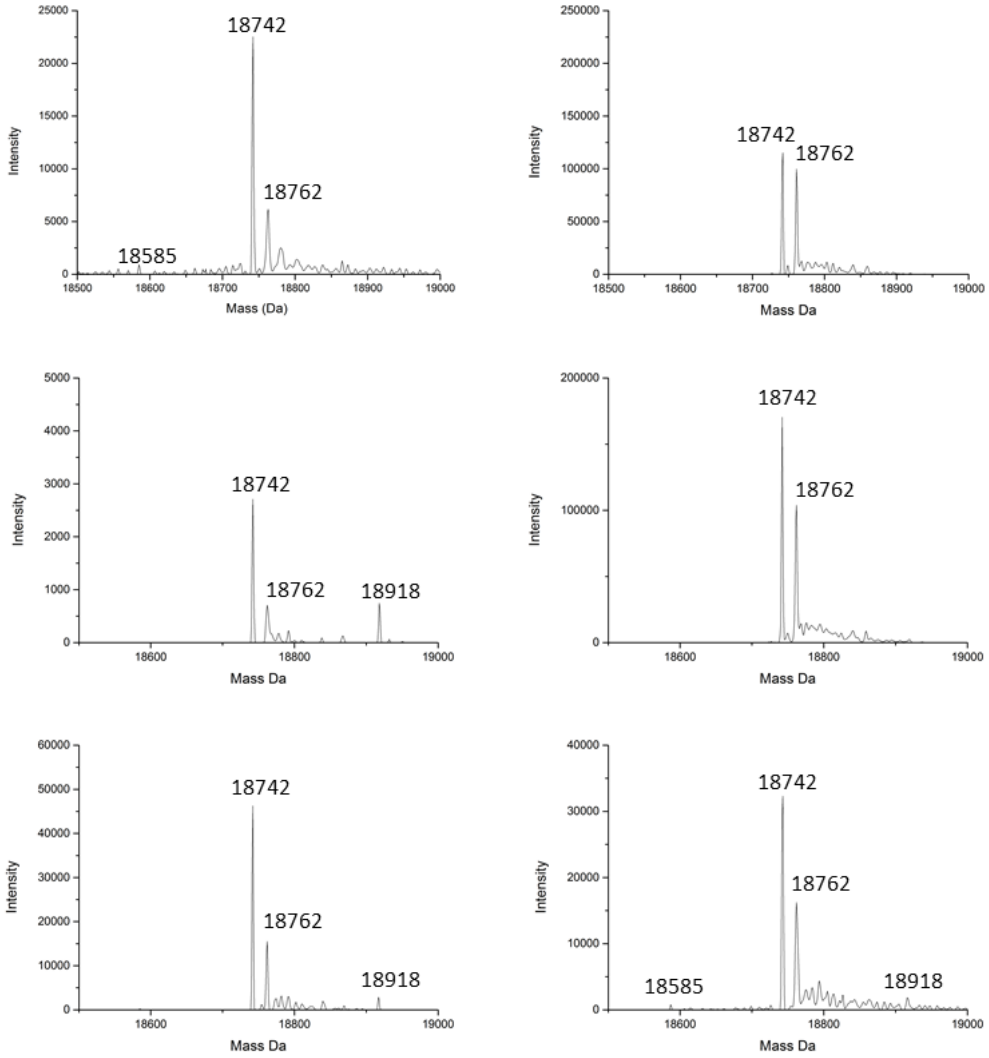


Figure 3.59 Summary of MUP expression patterns in male *M. spicilegus* urine. Accurate molecular weights of the MUPs in *M. spicilegus* urine from six different males was determined by ESI-MS. Urine samples were desalted as described in Chapter 2 and diluted 1:2 into 0.1% formic acid. Samples were injected onto a C4 desalting trap and the masses of the proteins of interest at approximately 20 kDa were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). All males were singly housed with no previous adult contact with other males/females when urine samples were recovered.

these MUPs were present in the 18585 Da MUP fraction, but sequencing the two purified MUPs first allowed definitive sequencing of the 18585 Da MUP. The use of two proteases provided overlapping peptide sequence information, and *de novo* sequencing was using PEAKS software. The peptide sequence data produced by the software was usually high quality with high confidence scores. The non-sex specific 18918 Da could not be purified by SAX chromatography due to its consistently low concentrations in both male and female urine. *De novo* sequencing of this MUP was by using low-complexity female urine samples, which was again digested using two proteases to produce overlapping sequence information. Unfortunately, only the majority of the 18918 Da MUP could be sequenced – no MS/MS spectra was observed for a small part of the protein sequence. The software uses a data refinement step which discards low quality spectra and poorly sequenced peptides, which may explain this. Despite manual searching of MS/MS spectra, no likely sequence was discovered. The fact that this MUP was not purified prior to *de novo* sequencing meant care had to be taken to not assign a sequence that 'fitted' into the rest of the protein sequence mass-wise, as it may have belonged to another protein present in the sample.

For the purposes of sequence alignment, the unsequenced part of the 18918 Da MUP was assumed to be the same as the sequence seen in *M. m domesticus* MUP 6, due to the high similarity between the sequences of the two. This resulted in the MUP sequence having a molecular mass 36 Da less than 18918 Da. An amino acid substitution gave the protein the correct molecular weight, and this was used for sequence alignment. The three male specific MUPs had sequences highly homologous to each other, and the non sex specific MUP shared much less homology with the three male specific MUPs. All *M. spicilegus* MUPs were aligned with all 21 *M. m. domesticus* MUPs and phylogenetic analysis was performed. It was confirmed that the three male specific *M. spicilegus* MUPs were very similar to central *M. m. domesticus* MUPs, with 18742 Da and 18585 Da MUPs being very similar to the male specific *M. m. domesticus* MUP 7 in particular. The 18762 Da MUP appears to be most similar to *M. m domesticus* MUPs 13 and 17. The non sex specific 18918 Da MUP is very similar to the peripheral *M. m. domesticus* MUP 6.

The presence of male specific MUPs in male urine, along with the variance in MUP expression seen in male urine and the similarity of these MUPs to *M. m. domesticus* MUPs, suggested that these MUPs may have a functional role in the unusual sexual and social behaviour observed in *M. spicilegus*, recorded in previous literature. Three

different experiments were set up at the Leahurst campus to determine whether contact between males and females, breeding between males and females and contact between males caused differences in their MUP expression output in the aim of determining a possible role of MUPs in their behaviour. Female mice did not alter their MUP expression upon contact with male mice, but male mice increased their MUP expression fairly significantly upon contact with females. Of particular interest was the apparent increase in male specific 18762 Da MUP expression in male urine upon contact with females, indicating that this MUP may play a role in the attraction of females. After male and female pairs were put together for breeding, all five males reduced their MUP expression, three of them significantly, including that of the 18762 Da MUP. Again, female MUP expression did not appear to change. In male/male pairs, the only pair of mice that appeared to significantly alter their MUP output during contact and after interactions were the mice that displayed significant aggressive behaviour during the interactions. One male in this pair, after interactions, significantly increased their MUP output; the other male significantly reduced their MUP output after interactions, along with their expression of the male-specific 18742 Da and 18762 Da MUPs, suggesting that increased MUP output (especially of these two male-specific MUPs) may also have a role in signalling dominance. More of these experiments need to be undertaken to definitively link male MUP expression and dominant behaviour. These experiments only gave a brief insight into the possible roles of *M. spicilegus* MUPs, in particular the male-specific MUPs, due to the relatively small number of animals sampled. There does, however, seem to be a fairly noticeable link between male MUP expression and contact with females, as well as male MUP output and aggressive behaviour amongst males.

Even though the male specific MUPs, which are those of particular interest in terms of possible roles in *M. spicilegus* behaviour, have been fully sequenced, it would be beneficial to obtain full accurate sequence information for the 18918 Da MUP. An alternative mass spectrometric method, such as electron transfer dissociation (ETD), may provide additional sequence coverage, or the development of a SAX chromatography method that provides better resolution of MUPs may allow the 18918 Da MUP to be purified from male urine, where MUP concentration is higher than in female urine. This would allow the possibility of producing recombinant versions of all four *M. spicilegus* MUPs to be used in further behavioural experiments, for example, are female mice attracted to the male specific MUPs? Are they more attracted to the 18762 Da MUP than the others? Do these MUPs promote aggressive behaviour in male mice? Does the 18918 Da MUP, seen in both sexes, elicit a

response in male or female mice? Answers to these questions would build significantly on the behavioural experiments conducted in this Chapter, providing a deeper insight into the unusual social and sexual behaviours this species has been found to exhibit. Further studies into the interactions of *M. spicilegus* MUPs with volatile pheromonal components in urine may also provide information on the roles of these MUPs in *M. spicilegus* behaviour.

Chapter 4: Communal nursing in the house mouse

4.1 Introduction

4.1.1 Communal nursing

Communal care, the sharing of parental responsibilities between multiple individuals, is common in many animals, including rodents (Gittleman 1985; Jennisons & MacDonald 1994; Emlen 1995; Hayes 2000). Communal nesting is a form of communal care where multiple females raise their offspring in the same nest, usually formed by groups of females residing and giving birth in one male's territory (Manning *et al*. 1995; Weber & Olsson 2008). Some communally nesting mammals also exhibit communal nursing, which is the sharing of milk with non-offspring (Packer *et al*. 1992). Female house mice communally nurse offspring in communal nests, in wild, semi-wild and laboratory environments (Wilkinson & Baker 1988; Konig 1994; Manning *et al.* 1995). This has led to many hypotheses being formed as to whether communal nursing is beneficial to mothers or offspring. Lactation involves high energetic costs and influences a mother's future reproduction, so it is unlikely that females provide milk to non-offspring without advantage (Konig *et al*. 1988; Hayes 2000).

Hayes (2000) reviewed hypotheses that have been formulated regarding the benefits and costs of communal nesting and nursing in rodents, including the likelihood that communal nursing increases reproductive success and provides fitness benefits, with increased pup survival, growth and weaning weights, group defence and thermoregulation (Gittleman 1985; Konig 1994). Hypotheses also suggest that communal nursing is a means by which females can be rid of excess milk that may cause discomfort. Despite this, along with the energetic cost of communal nursing, other possible disadvantages include food competition, increased visibility to predators and transmission of parasites. Other non-adaptive arguments include that nursing is indiscriminate, and that mothers are unable to prevent milk stealing (Konig 2006).

Konig (2006) was able to reject the hypothesis that milk stealing occurs, as non-offspring nursing is an integral part of female house mouse behaviour. Indiscriminate nursing can also be rejected, as female house mice are able to breed and nurse solitarily despite another female reproducing in the same territory. The study was

able to confirm that communal nursing increased reproductive success, whether or not the females were related. It was observed, however, that mutualistic cooperation was higher in nests where females were related. This indicates that kin recognition and selection were important during the evolution of communal nursing.

The fact that kin recognition appears to play a role in communal nursing provides further questions, especially since the hypotheses regarding indiscriminate nursing and milk stealing have been rejected. The benefits of communal nursing have been outlined, and while female house mice appear to take turns to feed the pups in the nest (Wilkinson and Baker, 1998; Konig 1989), it could be that females recognise their own young and invest in those more than non-offspring, and it could also be that females invest more in the offspring of their closer relatives (*i.e.* sisters) than their more distant relatives (*e.g.* cousins). It may be that females do not discriminate between their offspring and their non-offspring, and invest cooperatively more in larger, healthier pups.

Previous studies have been primarily based on the time a female spends retrieving pups, and on pup weaning weights. Analyses on communal nursing in BALB/c mice include age differences in litters nesting communally, survival rate of pups in communal nests, and the relatedness of communally nesting mothers. Manning *et al*. (1995) discovered that in communal litters where there was an age difference of 3-4 days, mothers of the older litters preferentially retrieved their own pups, but the mothers of younger litters did not discriminate between their own offspring and alien offspring. In the same paper, Manning describes how pup survival is greater in communal nests, with the probability of pup survival to weaning over double that of in single nests (69% compared to 33%). Konig (1994) stated that under natural conditions, females sharing a nest are often related and their pups typically have the same father, meaning it his highly likely that females nurse closely related offspring when nesting with a familiar female. Whilst familiar females (regardless of relatedness) nesting communally improved female reproduction and pup survival, female groups which were familiar and related also produced the largest litters with heaviest pups, suggesting that relatedness of females results in increased investment in litters.

4.1.2 Metabolic stable isotope labelling

Since the work of Schoenheimer *et al.* (1938), it has been realised that proteins in a cell are in a dynamic state of turnover, and whilst proteins in a cell or tissue can be synthesised at a similar rate, differences in their rate of degradation result in differences in steady–state concentrations of these proteins (Claydon *et al*. 2012). Simple proteome turnover studies in isolated cells often take place in minimal media with the introduction of labelled amino acids (for example, [$^{13}C_6$] arginine), allowing the labelling of all proteins containing that amino acid residue in a sample. The size of the labelled amino acid pool is substantially larger than the intracellular pool, meaning proteins are rapidly fully labelled (Claydon *et al*. 2012). In more complex systems, such as animals, protein turnover studies using stable isotope labelling is more difficult: incorporation of the labelled amino acid cannot be instantaneous due to the biomass of the animal tissues contributing to the precursor pool, and the introduction of labelled amino acid is best done via diet (causing fluctuations in label incorporation due to the active/inactive periods spent by an animal). Different tissues are affected differently, in terms of label incorporation, by feeding fluctuations; the liver and small intestine are most exposed to labelling variation, whereas tissues such as muscle are more metabolically remote and so are less affected by feeding fluctuations (Claydon *et al*. 2012). Claydon *et al.* (2012) designed a simple experiment to incorporate labelled amino acids into mice via a 50% labelled (therefore palatable to the mice, since they find a fully synthetic diet unpalatable) rodent diet, using MUPs in urine to monitor the labelling of the liver - after synthesis in the liver, MUPs are transported to the kidneys through the bloodstream and excreted in urine (Robertson *et al*. 1998; Beynon and Hurst 2003; Hurst and Beynon 2004). There is no evidence of MUP uptake during renal filtration, and so MUPs reflect the properties of the precursor RIA in the liver (Armstrong *et al*. 2005). This non-invasive sampling method would allow the kinetics of the labelling of the liver precursor pool to be evaluated.

This relatively inexpensive and simple experimental technique used for assessing the turnover of proteins in mice could, in principle, be used in determining the proportion of lactation investment from communally nursing female mice – female mice could be fed a labelled diet, and the kinetics of labelling could be determined by mass spectrometric analysis of MUPs in urine. These mice, feeding their pups, should pass their labelled proteins onto their pups, which then allows the label to be incorporated and tracked in their offspring. Analysis of stomach contents and tissue samples from

pups, nursed by the mice who are fed a labelled diet, will determine whether the label is being passed on from the mothers and being incorporated into the pup precursor pool.

4.1.3 Aims

Whole animal metabolic labelling with stable isotope labelled amino acids can be used to determine the proportion of lactation investment from communally nursing female house mice and to ascertain any discriminatory factors. In this study, the aim was to introduce a suitably labelled amino acid into the animals via a semi-synthetic diet until whole animal metabolic labelling had been reached, and these labelled amino acids would be passed on to the offspring through feeding, determining the proportion of investment from each mother in the nest. An experiment was set up where two female mice nest communally with their new-born pups, and each mother has access to a semi-synthetic isotope - labelled diet. These diets are differently labelled, in order to track the different labels in their offspring, and each mother only has access to one of these diets. Heavy and light isotopes of the amino acid are required as this causes the resultant heavy/light peptides from the proteins expressed by the animal to behave the same in mass spectrometry analysis, yet provide a reasonable mass difference to differentiate between isotopes for calculation of label incorporation. Labelled amino acid incorporation can be assessed in the female mice (mothers) by mass spectrometric analysis of MUPs expressed in urine, and analysis of the stomach contents of their pups (containing the milk and resultant labelled amino acids passed on from the mothers) can be used to ensure the labels have been passed on. Analysis of other pup samples (liver, muscle, urine) can be used to determine the proportion of lactation investment from mothers over a period of time. Label incorporation, and therefore investment, are observed in litters of pups where their mothers are related and in litters where their mothers are unrelated, to assess whether any discrimination in investment is evident in relation to relatedness of the female pairs in the nest. Weight of pups, litter age differences and litter sizes are also considered when assessing label incorporation, aiming to identify whether these factors affect the proportion of investment received.

4.2 Results and discussion

4.2.1 Successful incorporation of labels into mouse diet

4.2.1.1 Diet

To ensure the diets provided to the mice were 50% labelled with [$^{2}H_{8}$] valine or [$^{13}C_{6}$] lysine, two unlabelled diets (of Certified Rodent Diet 5002) were separately prepared, having obtained the percentage of valine and lysine in the diet. In one diet, [$^{2}H_{8}$] valine was added in an amount to achieve a percentage of 50% labelled valine, and in the other diet, [$^{13}C_{6}$] lysine was added in an amount to achieve 50% labelled lysine. The diet was prepared as described in Chapter 2 and the Supplementary Material, and fed to mice by Dr J.P. Green at the Mammalian Behaviour and Evolution Group, Leahurst Campus, University of Liverpool.

The initial preparatory experiment for the lactation investment study was to ensure the semi-synthetic heavy labelled diets were palatable to adult female mice, in order to ensure rapid and complete whole animal metabolic labelling. Four female BALB/c mice were fed either a [$^{13}C_{6}$] (heavy) lysine or [$^{2}H_{8}$] (heavy) valine labelled diet, with the label incorporated so that the precursor relative isotopic abundance (RIA) should be approximately 0.5, over a period of five days. Precursor RIA values are calculated by

$$RIA = I_H / (I_H + I_L)$$

where I_H is the signal intensity displayed by the heavy labelled peptide (calculated by summing the intensities of the monoisotopic (M_0), M_0 + ^{13}C (M_1) and M_0 + 2 ^{13}C (M_2) isotopic peaks), and I_L is the signal intensity of the light (unlabelled) peptide, calculated in the same way. The incorporation of the labels into the mice result in all proteins in the animal to be 50% either heavy valine or heavy lysine labelled; the rate at which a protein reaches a labelling maximum depends on that protein's rate of turnover. The most simple and effective way of assessing label incorporation was through the analysis of MUPs present in urine; the fact that MUPs are quantitatively excreted in urine would determine whether the diet was being consumed by the mice, and how long it would take for labelling of the liver precursor pool to reach a maximum. Calculating RIA values in MUPs do not allow determination of protein turnover rates as these proteins are simply excreted from the animal, and so do not degrade, but these values do allow analysis of the liver precursor pool (Claydon *et al*. 2012). The

rate at which both semi-synthetic diets were consumed by the mice is indicative of the diets being found palatable by the mice (Figure 4.1), and should ensure rapid labelling of the precursor pool. As the mice being fed the diet were fully-grown adults, their rate of growth is negligible and so this did not need to be factored in when calculating the RIA in urinary proteins.

4.2.1.2 Assessing the rate of label incorporation

The urine was digested in-solution and analysed using MALDI-ToF-MS as described in Chapter 2. The MUP peptides in the MALDI-ToF-MS spectra of the digested mouse urine (Figure 4.2) were identified by searching the MUPs_mature database via the Mascot search engine. Analysing the sequences of each identified MUP peptide allowed the identification of valine-containing peptides – all peptides contained one lysine residue as endoprotease Lys-C was used as the protease, resulting in all peptides being lysine-terminated. Lys-C was chosen as the most suitable endoprotease as it gives better peptide coverage of MUPs for MALDI-ToF-MS analysis (Beynon *et al*. 2002). This also ensured that each peptide generated from the proteolysis of MUPs contained a lysine residue, resulting in more [$^{13}C_6$] lysine labelled peptides for calculating precursor RIAs.

Four valine-containing peptides were identified in the spectra, two of which were di-valine peptides (contained two valine residues). From the 'light' (unlabelled) peptide masses and the sequence information, the masses of the corresponding 'heavy' lysine and valine peptides were calculated. For all lysine-containing peptides, the mass of the corresponding 'heavy' labelled peptide is 6 Da heavier than the 'light' unlabelled peptide, due to the incorporation of six ^{13}C into the lysine residue. For the two single valine-containing peptides, the mass of the corresponding 'heavy' peptide should be 8 Da heavier due to the incorporation of eight deuterium atoms into the valine residue. For the two di-valine peptides, a fully labelled peptide should be 16 Da heavier than the 'light' unlabelled peptide, as eight deuterium atoms are incorporated into each of the valine residues (Table 4.1).

4.2.1.3 Transamination of [$^2H_8$] valine

During analysis of the mono-valine containing peptides seen in the mass spectrometric analysis, it was noted that the corresponding heavy labelled peptide was only +7 Da heavier, rather than the expected +8 Da shift. In di–valine peptides, a mass shift of +14 Da was seen, rather than the expected +16 Da. In both instances, the labelled peptide isotopologue profiles were different to that of the unlabelled

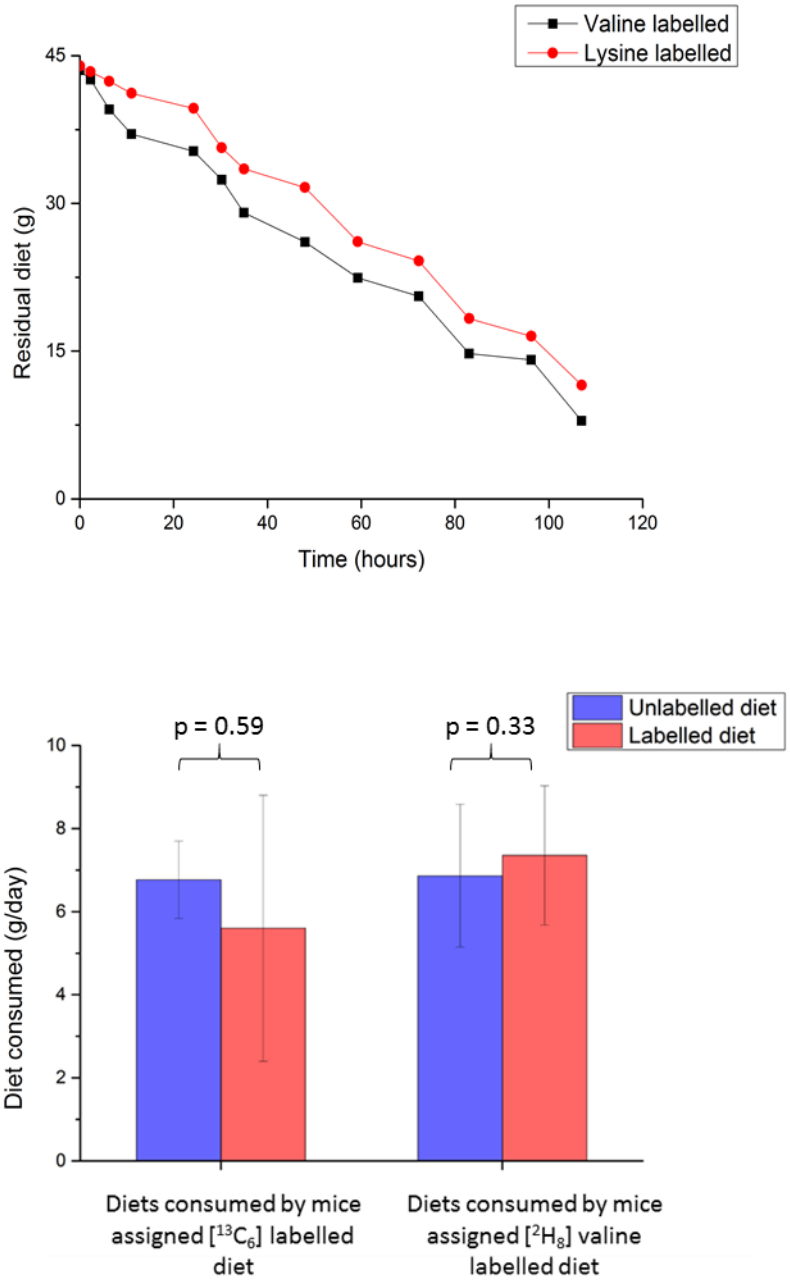


Figure 4.1 The consumption of the two different labelled diets.
Top: The average amount of unlabelled diet consumed by the four mice (two assigned the [$^{13}C_6$] lysine labelled diet (left) and two assigned the [$^{2}H_8$] valine labelled diet (right)), over the seven day period prior to the introduction of the labelled diets, are shown in blue (error ± SD, n=7); the average amount of labelled diets consumed by the mice during the five day experiment are shown in red (error ± SD, n = 5). There was no significant difference in the amount of unlabelled/ [$^{13}C_6$] lysine labelled diet consumed (Welch two sample t-test: t = 0.62, df = 2.14, p = 0.59) or in unlabelled/ [$^{2}H_8$] valine labelled diet (Welch two sample t-test: t = 1.06, df = 4.99, p = 0.33). Bottom: The residual [$^{2}H_8$] valine labelled diet and [$^{13}C_6$] labelled diet, each accessed by two mice, was weighed at various intervals during the course of the five day experiment.

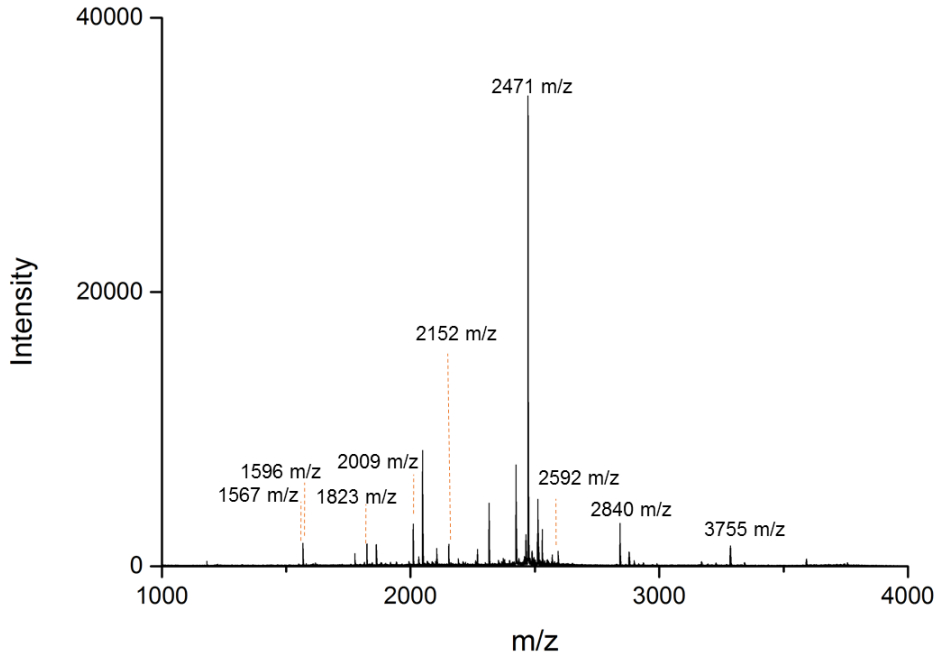


Figure 4.2 Peptide mass fingerprint of female BALB/c mouse urine prior to labelled diet.
Whole urine was digested using the in-solution digestion protocol, with Lys-C as the protease. 1 µl of digested sample was mixed 1:1 with matrix solution before being spotted onto the MALDI target plate. Mass spectra were acquired using the method described in Chapter 2. Labelled are the MUP peptides identified by database search via Mascot search engine.

Table 4.1 Summary of the MUP peptides identified in MALDI-ToF-MS analysis of female BALB/c urine.
Included are the mass of the identified MUP peptides, the MUPs in which these peptides are found, the sequences of the peptides and the masses observed if the peptides are labelled with [$^{2}H_{8}$] valine or [$^{13}C_{6}$] lysine. All peptides are lysine terminated due to the use of Lys-C as the protease, so if a mouse was fed a [$^{13}C_{6}$] lysine diet, all observed peptides would be labelled. Four of the observed peptides contain a valine residue, so if a mouse was fed a [$^{2}H_{8}$] valine diet, only these MUP peptides would be labelled.

MUP peptide (Da)	MUP(s)	Sequence	[$^{2}H_{8}$] valine labelled peptide (Da)	[$^{13}C_{6}$] lysine labelled peptide (Da)
1567.7	1,7,8,9,10,11,12,13,14,16,17,18,19	EASSTGRNFNVEK	1575.7	1573.7
1596.8	1,2,7,8,9,10,11,12,13,14,15,16,17,18,19	INGEWHTIILASDK	-	1602.8
1823.8	1,2,7,8,9,10,11,12,13,14,15,16,17,18,19	TDYDNFLMAHLINEK	-	1829.8
2009.9	1,2,7,8,9,10,11,12,14,15,16,17,18,19	AGEYSVTYDGFNTFTIPK	2017.9	2015.9
2152.9	1,2,9,10,11,12,14,15,16,18,19	FHTVRDEECSELSMVADK	2168.9	2158.9
2471.1	1,2,7,8,9,10,11,12,13,14,15,16,17,18,19	DGETFQLMGLYGREPDLSSDIK	-	2477.1
2592.3	1,7,12,15	HGILRENIIDLSNANRCLQARE	-	2598.3
2840.5	1,2,7,9,11,15,16,18,19	IEDNGNFRLFLEQIHVLENSLVLK	2856.5	2846.5
3754.8	2,8,9,10,11,13,14,16,17,18,19	ERFAQLCEEHGILRENIIDLSNANRCLQARE	-	3760.8

peptides. The reason for the 7 Da shift (in mono-valine peptides) and the 14 Da shift (in di-valine peptides) is because the deuterium atom present on the α-carbon atom of [$^{2}H_{8}$] valine is metabolically labile and is known to transaminate *in vivo* (Doherty *et al*. 2005). The reason for the different isotopologue profiles seen in labelled peptides is the overlap of the [$^{2}H_{7}$] valine (+7 Da) and the [$^{2}H_{8}$] valine (+8 Da) label incorporation. If the +7 Da and +8 Da peptides were observed alone, they would have an isotopologue profile identical to the unlabelled peptide. The overlap causes the M_0 (^{12}C) peak of the +7 Da peptide to appear smaller than expected, with the +8 Da peptide M_0 peak adding to the intensity of the +7 Da M_1 (^{13}C) peak, making the M_1 peak more intense. In di-valine peptides, the +14 Da shift is due to the incorporation of transaminated ([$^{2}H_{7}$]) valine into both of the valine residues; the +15 Da shift is due to the incorporation of [$^{2}H_{7}$] valine into one of the valine residues, and the incorporation of [$^{2}H_{8}$] into the other. The +16 Da shift is due to the incorporation of [$^{2}H_{8}$] valine into both valine residues (Figure 4.3). The overlap of these three observed peptide shifts causes the different isotopologue profile seen in the doubly-labelled valine containing peptides.

4.2.1.4 Calculation of precursor RIA

The precursor RIA values were calculated for the two most abundant MUP peptides for the mice fed a [$^{13}C_{6}$] lysine labelled diet, and for the two single valine-containing MUP peptides for the mice fed a [$^{2}H_{8}$] valine labelled diet. Each of the four mice were sampled once prior to being fed the labelled diet, and for days 1 and 2 of labelling, all four animals were sampled in the morning, afternoon and evening. For days 3-5 of labelling, all animals were sampled morning and evening. MS isotope (UCSF) (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msisotope) was used to calculate the isotopologue profiles for each unlabelled MUP peptide used for the calculations, to confirm that the unlabelled and corresponding labelled peptides were suitably separated and that the intensities of unlabelled peptide peaks did not significantly contaminate the intensities of the labelled peptide peaks. For example, for valine containing MUP peptide 2009 Da (AGEYSVTYDGFNTFTIPK), MS isotope determined that the +7 Da labelled peptide would only contain 0.05% of the unlabelled peptide peak intensity, and that the +8 Da labelled peptide would contain only 0.01% of the unlabelled peptide peak intensity. For the lysine containing MUP peptide 2471 Da (DGETFQLMGLYGREPDLSSDIK), the labelled peptide (+6 Da shift) would contain only 0.68% of the unlabelled peptide peak intensity. Although the peak

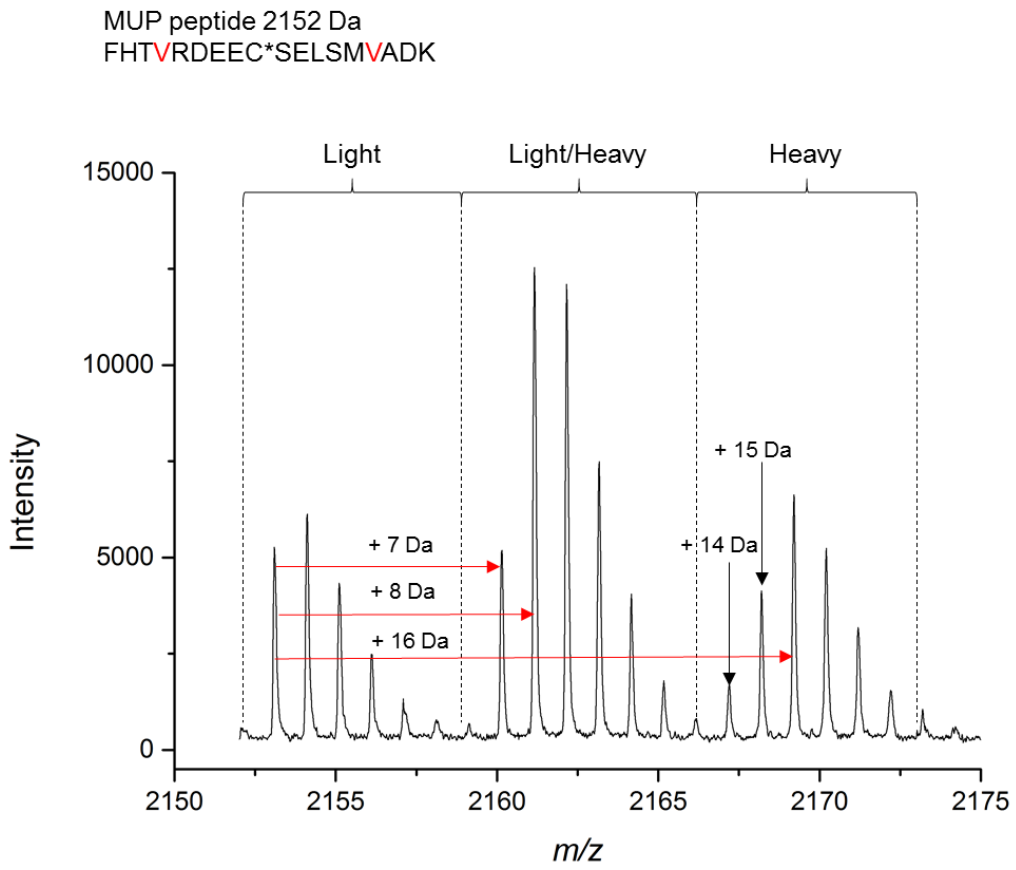


Figure 4.3 Light, light/heavy and heavy isotope profile of a MUP peptide containing two valine residues, from the urine of a mouse fed [$^2H_8$] valine labelled diet.
The MALDI-ToF-MS spectrum obtained for the di-valine MUP peptide shows the presence of the unlabelled ('light') peptide, the peptide where one of the two valines is labelled ('light/heavy'), and the fully labelled ('heavy') peptide. The + 7 Da shift (rather than + 8 Da expected from the labelling of a valine residue) is due to the transamination of the [$^2H_8$] valine, resulting into the loss of the alpha-carbon deuteron which is replaced by a hydrogen**.** The + 8 Da shift is the labelling of one of the valine residues in the peptide, and the + 16 Da shift signifies the labelling of both valine residues in the peptide, with the + 14 Da and + 15 Da shifts due to the transamination of [$^2H_8$] valine.

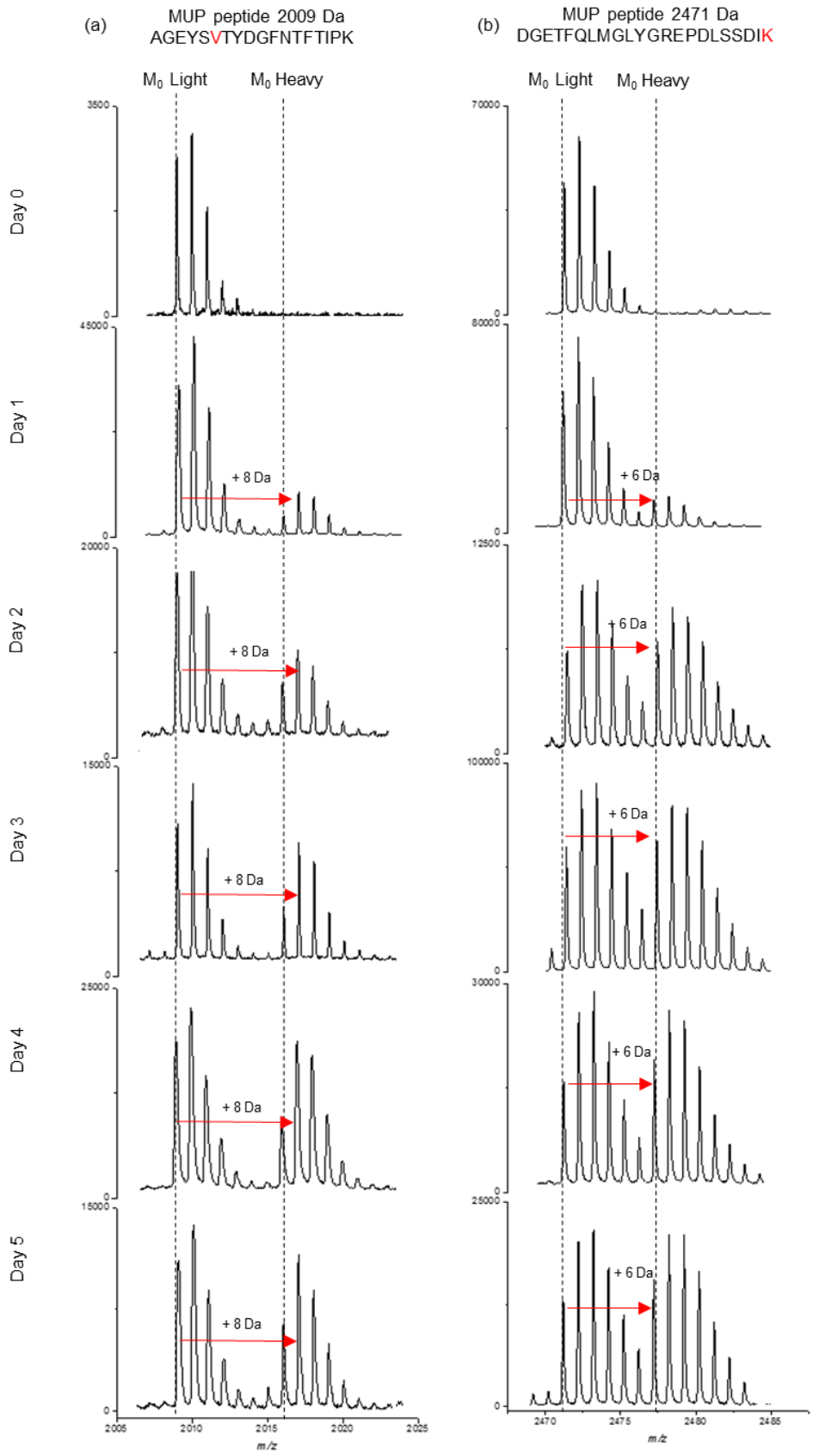
contamination is slightly higher for the lysine containing peptides, it is still negligible and will not significantly affect the RIA calculations.

The unlabelled and corresponding labelled peptide profiles for these two MUP peptides, from urine sampled from mice on the morning of day 0 of being fed a labelled diet, and urine sampled in the evenings of days 1 – 5 of the mice being fed a labelled diet, are shown in Figure 4.4. Figure 4.4 (a) shows the course of how valine containing MUP peptide 2009 Da labels over the course of the 6-day experiment, obtained from a mouse fed the [$^{2}H_{8}$] valine diet, and Figure 4.4 (b) shows the same but for lysine containing MUP peptide 2471 Da, from the urine of a mouse fed the [$^{13}C_{6}$] lysine diet. In both peptides, it was noted that both the labelled and unlabelled versions showed a different isotopologue profile than would be normally expected, where M_0 peak is the most intense, with M_1 peak slightly less intense and M_2 peak less intense again. This 'normal' peptide profile was seen in the smaller MUP peptides 1568 Da and 1597 Da, and this profile was confirmed using MS isotope (UCSF). When MS isotope was used to analyse the profiles of the larger MUP peptides, it confirmed that the isotopologue profiles seen in the MALD-ToF analysis (where the ^{13}C (M_1) peak is more intense than ^{12}C (M_0) peak) is the peptide profile expected. This is because in larger peptides, there is a greater natural abundance of ^{13}C present than in smaller peptides.

From the peptide profiles in Figure 4.4, no corresponding labelled proteins are seen on day 0, before the mice were fed the labelled diets, as would be expected. Over the course of the five labelling days, the abundance of labelled peptide increases each day, indicating that the mice are consuming the diets and that both labels are incorporating into the mice assigned either the [$^{2}H_{8}$] valine or [$^{13}C_{6}$] lysine labelled diet. To determine the point at which the precursor pool of the liver had reached its labelling plateau, the urine samples from each of the four mice (two of which were fed the [$^{2}H_{8}$] valine diet, the other two fed the [$^{13}C_{6}$] lysine diet) were collected on the morning of day 0, the morning, afternoon and evening of days 1 and 2 of labelling, and in the morning and evening of days 3, 4 and 5 of labelling. No samples were taken at night time as this is the time that the mice are inactive (as this is their 'day' period).

Each urine sample collected was subjected to in-solution digestion with Lys-C and analysed using MALDI-ToF-MS. For each of the samples from the two mice fed the [$^{2}H_{8}$] valine diet, the two mono-valine containing MUP peptides were identified in the MALDI-ToF-MS spectra, and for each peptide, the intensities of the unlabelled and

Figure 4.4 Incorporation of heavy labelled amino acids over the course of 6 days, shown by the relative abundances of heavy/light peptide profiles in urine from two BALB/c females, (a) fed a diet labelled with [$^{2}H_{8}$] valine and (b) fed a diet labelled with [$^{13}C_{6}$] lysine.
Urine samples were taken from the females at intervals over a six-day period. Shown are the unlabelled ('light') and labelled ('heavy') MUP peptide profiles taken from urine collected in the morning (Day 0) and evenings of each day for each mouse. The 'heavy' valine peptide is only 7 Da larger than the light peptide, though the heavy valine label involves the incorporation of 8 deuterium atoms in each valine, due to the transamination of [$^{2}H_{8}$] valine. The 'heavy' lysine MUP peptide is 6 Da heavier than the light MUP peptide, due to the incorporation of six ^{13}C into the lysine residue.


(a)
MUP peptide 2009 Da
AGEYSVTYDGFNTFTIPK
(b)
MUP peptide 2471 Da
DGETFQLMGLYGREPDLSSDIK
M0 Light
M0 Heavy
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5
+ 8 Da
+ 6 Da
m/z

labelled peptides were recorded and the RIA for that peptide was calculated using the formula stated previously. From this, an average RIA for each sample at each sampling time point (Figure 4.5). The same method of calculating RIAs was used for the samples collected from mice being fed the [$^{13}C_6$] lysine diet, except any two MUP peptides could be selected for the analysis due to the use of Lys-C for proteolysis, resulting in all peptides being lysine terminated (Figure 4.5).

Figure 4.5 shows how the calculated RIAs for samples collected in the morning are consistently lower than those calculated from the afternoon and evening samples. This fluctuation in RIAs is due to the mice being inactive during their 'day' period overnight, meaning no (or very little) diet will be consumed during this time. This means the degradation of existing unlabelled proteins over this time will be returning unlabelled valine/lysine residues to the precursor pool, without any contribution of labelled residues from the diet, diluting the proportion of [$^{2}H_8$] valine/[$^{13}C_6$] lysine present in the precursor pool and therefore reducing the calculated RIAs. Over time, this fluctuation appears to reduce, and this will be due to the fact that some degraded proteins will contain the label, returning both labelled and unlabelled residues to the precursor pool, reducing the dilution of the label in the pool. Daily RIA averages for each animal were then calculated to determine the time taken for each animal to appear to have the label fully incorporated (Figure 4.6). A precursor RIA value of around 0.5 indicated that the label is fully incorporated into the animal (as the diet is 50% labelled). Three of the mice (two fed [$^{2}H_8$] valine diet, one fed [$^{13}C_6$] lysine diet) reached a maximum RIA of 0.43, and the fourth mouse (fed [$^{13}C_6$] lysine) reached a maximum RIA of 0.45. These maximum RIAs were calculated from the average RIA calculated on the final day of the experiment for each mouse. The time taken for each mouse to have the label fully incorporated varied – for mouse 10903 (fed [$^{2}H_8$] valine), it appeared to take between 1 – 2 days for the precursor pool to become fully labelled, but for mouse 10904 (also fed [$^{2}H_8$] valine), it seemed to take between 3 - 4 days. For the two mice fed [$^{13}C_6$] lysine, mouse 10910 was fully labelled by day 2 of labelling, but for mouse 10911, labelling of the precursor pool took slightly longer at 3 days. This difference is likely to be due to the difference in the amount of diet consumed by the mice, but the experiment confirmed that the mice found the semi-synthetic labelled diets palatable and that full labelling of the protein precursor pool had happened by 3 – 4 days after the mice had been provided with the labelled diets. At this point, all proteins being synthesised in the liver and all proteins degrading contain both labelled and unlabelled valine/lysine, returning an almost equal amount of these to the precursor pool, meaning that the labels are fully incorporated into the mice.

Figure 4.5 The rate of heavy amino acid incorporation over 6 days in two mice fed [$^{2}H_{8}$] valine labelled diet and two mice fed [$^{13}C_{6}$] lysine labelled diet.
Urine samples were taken from the four BALB/c females at various time intervals over the six-day period, subjected to in-solution digestion with Lys-C and analysed using MALDI-ToF-MS. Red points are morning samples, blue are afternoon samples and black are evening samples. The average RIA was calculated from the three most abundant valine-containing (a) and lysine-containing (b) MUP peptides in each sample. Average maximum RIAs reached for each mouse are shown by a red line.

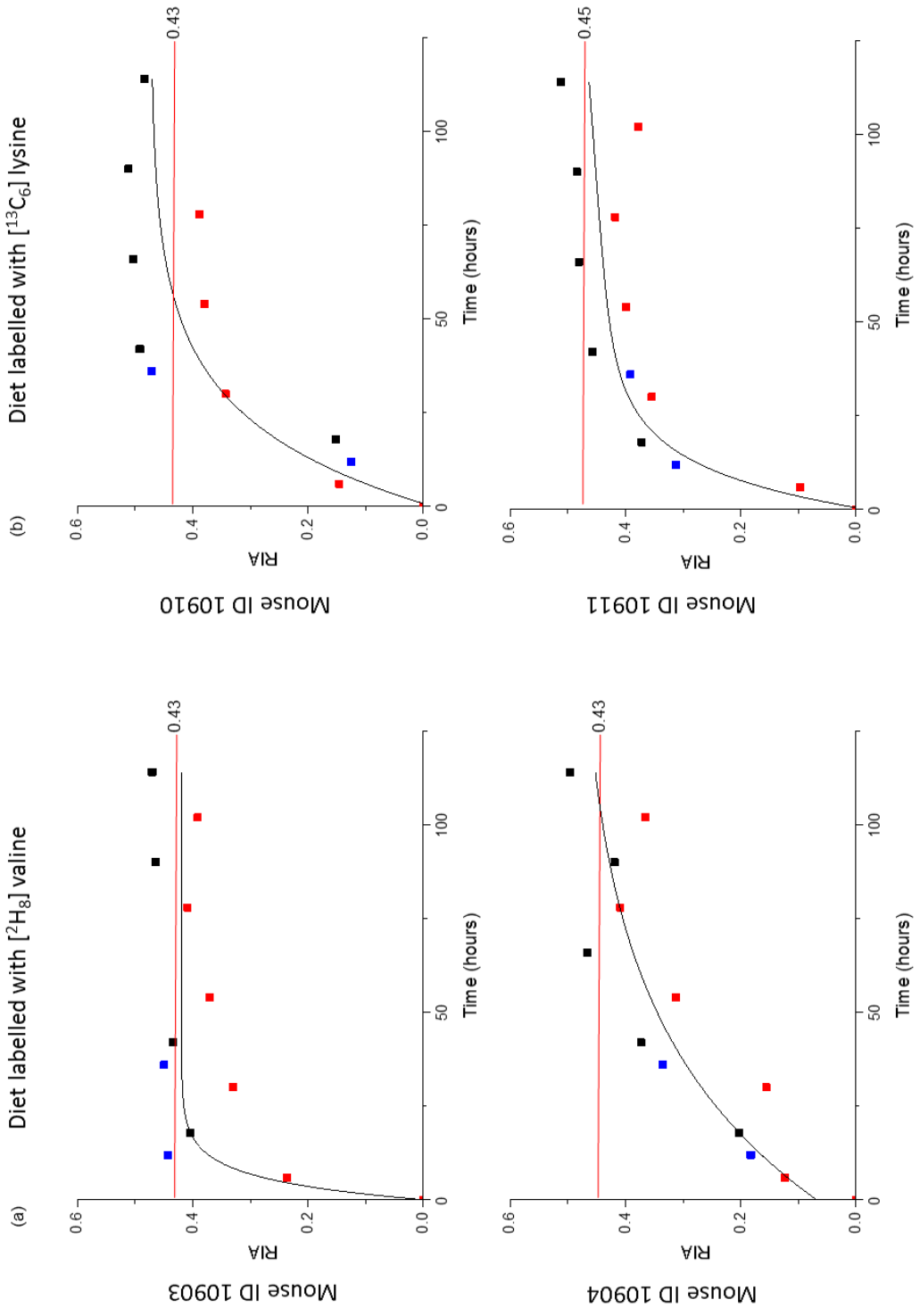
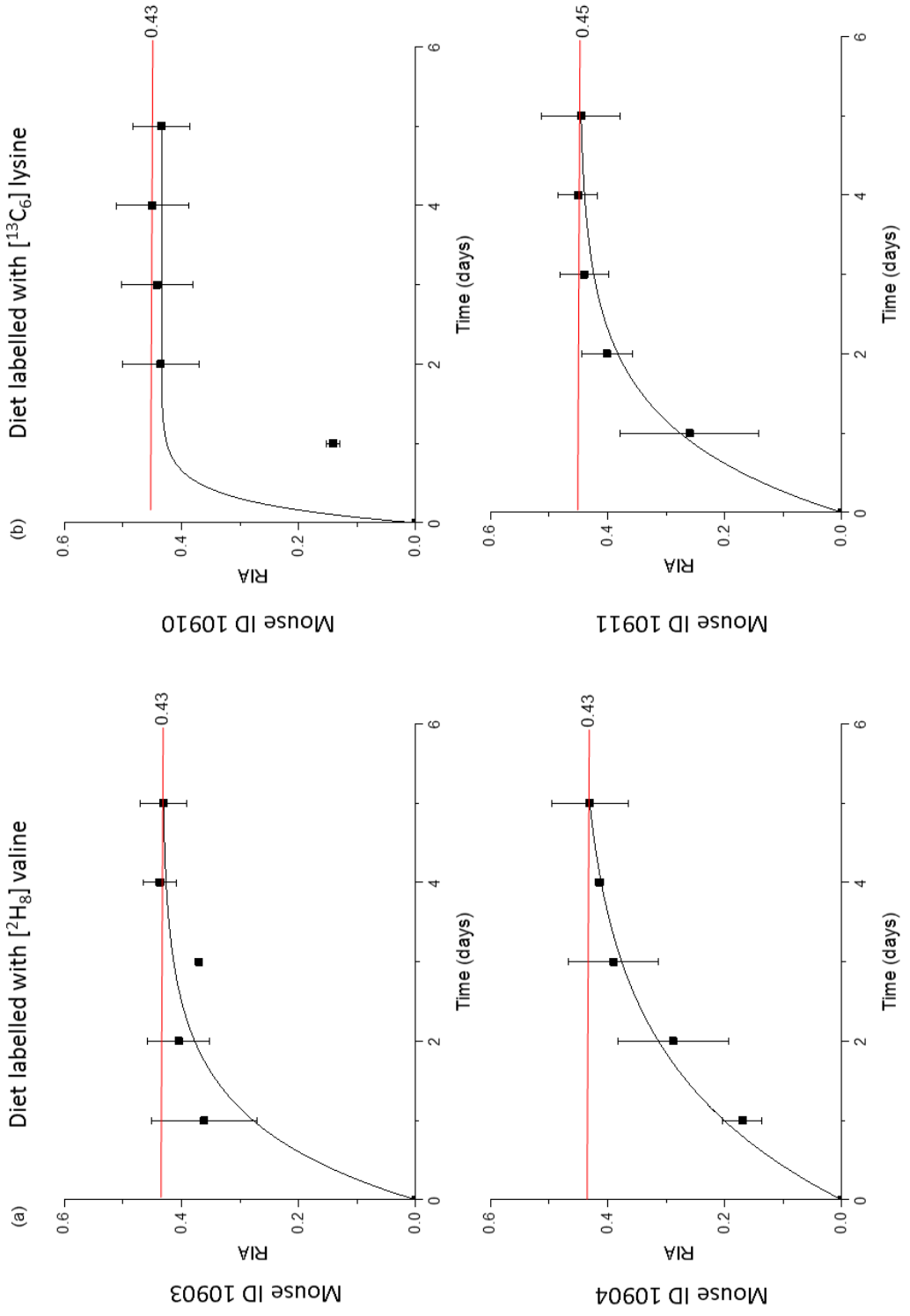

(a)
Diet labelled with [2H8] valine
(b)
Diet labelled with [13C6] lysine
Mouse ID 10903
Mouse ID 10904
Mouse ID 10910
Mouse ID 10911
RIA
Time (hours)
0.43
0.45
0.0
0.2
0.4
0.6
0
50
100

Figure 4.6 The rate of heavy amino acid incorporation over 6 days in two mice fed [$^{2}H_{8}$] valine labelled diet and two mice fed [$^{13}C_{6}$] lysine labelled diet.
Urine samples were taken from the four BALB/c females at various time intervals over the six-day period, subjected to in-solution digestion with Lys-C and analysed using MALDI-ToF-MS. An average RIA was calculated from the three most abundant valine-containing (a) and lysine-containing (b) peptides for each day, with error bars showing the standard deviation calculated for each daily RIA. Average maximum RIAs reached for each mouse are shown by a red line.


(a)
Diet labelled with [$^{2}H_{8}$] valine
(b)
Diet labelled with [$^{13}C_{6}$] lysine
Mouse ID 10903
Mouse ID 10904
Mouse ID 10910
Mouse ID 10911
RIA
Time (days)
0.43
0.45
0.0
0.2
0.4
0.6
0
2
4
6

4.2.2 Proof-of-principle

4.2.2.1 Feeders analysis

Having confirmed that the method of labelling mice worked successfully, the next step was to find a way to have two mice living in the same cage consuming two differently labelled diets, without having access to the other diet. This is so two female mice are able to live together and communally nurse their litters, with their investment in the pups tracked via the transfer of two different labels from the mothers to the pups. It is essential that the mothers and their litters can live together, but also that the mothers have access to only the diet to which they have been assigned. If they were to consume the labelled diet assigned to the other mouse, they would be 'contaminated' with the other label, and so tracking the investment from both mothers in their pups would be inaccurate. Professor J.L. Hurst developed a cage that would allow a pair of female mice to live and nest together, having unrestricted access to their pups, whilst having access to only the labelled diets each individual was assigned to. Figure 4.7 shows the set-up of the housing, where two female mice are able to share the same cage space, water, nest box and access to their respective diets. These diets are contained in feeding traps and are positioned at either end of the cage. Each female mouse has a subcutaneous RFID chip, and this chip can be read by the feeding mechanism. If, for example, a mouse assigned a [$^{2}H_{8}$] valine labelled diet approaches the correct diet, the feed will remain within the reach of the mouse. If, however, this mouse was to approach the [$^{13}C_{6}$] lysine labelled diet, the feeding mechanism would sense this via the chip carried by the mouse, and would subsequently move the diet away from the cage, out of the reach of the mouse. To determine whether this mechanism worked successfully, and whether this housing set-up would be suitable in allowing pairs of female mice to nest communally yet consume differently labelled diets, three pairs of female mice were chipped and housed in this situation for five days. In each of the three cages, the diets placed in one of the feeding traps was [$^{2}H_{8}$] valine, and in the other was [$^{13}C_{6}$] lysine. In each of the female pairs, one mouse was assigned the [$^{2}H_{8}$] valine diet, and the other the [$^{13}C_{6}$] lysine diet. Urine was sampled from each mouse daily over the course of the experiment, for the analysis of label in MUPs.

4.2.2.2 Analysis of MUPs for labelling

All urine samples were digested in-solution with Lys-C and subjected to MALDI-ToF-MS analysis. In each spectrum, the MUP peptides were analysed for the presence of the 'correct' label (the label arising from the labelled diet which the mouse was

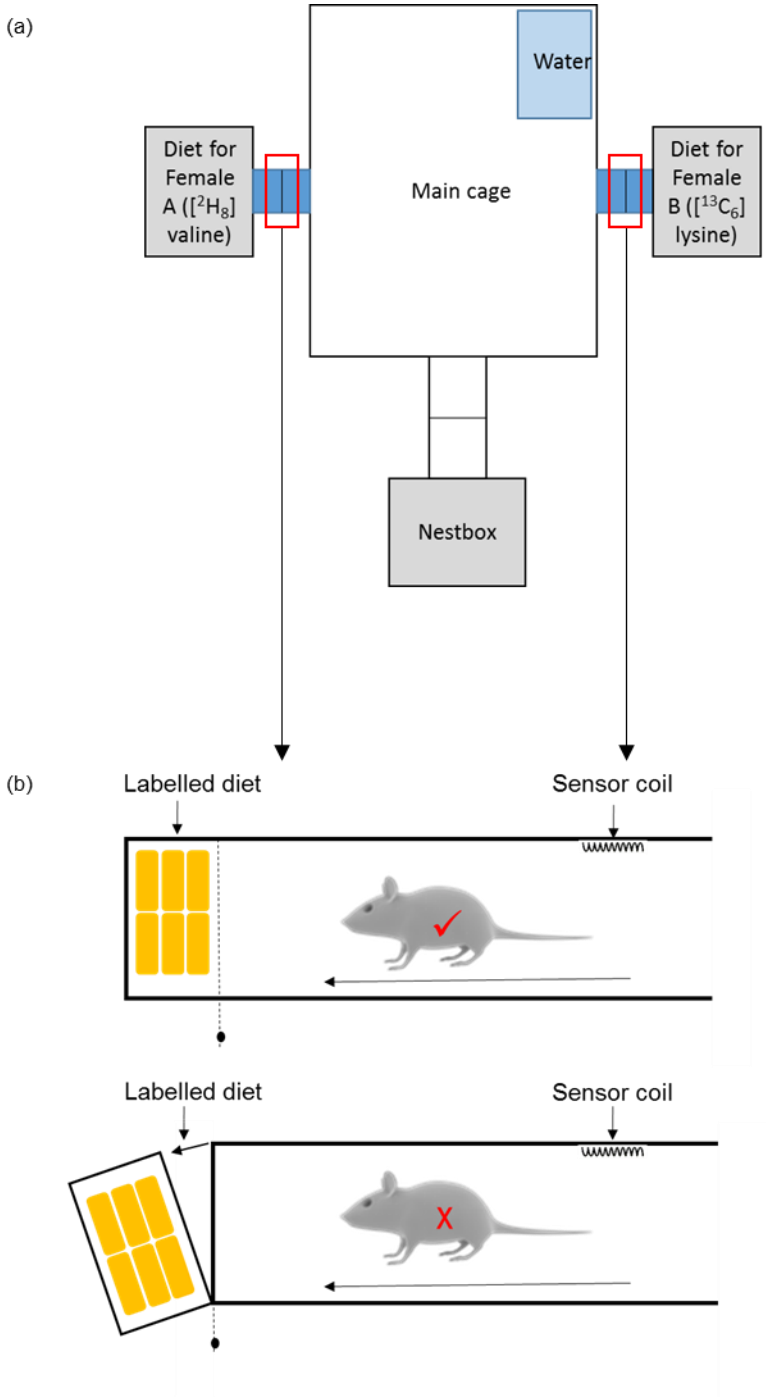


Figure 4.7 Diagram showing the feeding mechanism developed at Leahurst.
(a) Two feeding traps, each containing a different diet, are positioned at each end of the cage, with both mice having access to all parts of the cage. (b) Each female mouse has an RFID chip implanted in their necks, and this chip can be read by the sensor coil placed at the entrance of each feeding mechanism. If a mouse assigned a [$^{2}H_{8}$] valine labelled diet approaches the correct diet, the feed will remain in reach of the mouse. If the mouse was to approach the [$^{13}C_{6}$] lysine labelled diet, the feeding mechanism would sense this via the chip carried by the mouse, and would subsequently move the diet away from the cage, out of reach of the mouse. The data from these sensors are logged, as is the amount of time spent in the nestbox by each mouse.

assigned) to ensure the mechanism allowed the mouse to access the diet, and also the presence of the 'wrong' label (the label from the diet which the other mouse in the pair was assigned to) to ensure that the feeding trap moved away when the 'wrong' mouse approached it. With the two different labelled diets being [$^{2}H_{8}$] valine and [$^{13}C_{6}$] lysine, there is only a mass difference of 2 Da if both labels were to be observed in a single urine sample (the presence of [$^{2}H_{8}$] valine in the mouse gives a theoretical shift of +8 Da of the unlabelled MUP peptide, and the presence of [$^{13}C_{6}$] lysine gives a shift of +6 Da). Taking into account the transamination of valine, resulting in just a +7 Da shift of the unlabelled peptide, the identification of either label in the mass spectra must be done very carefully. It is easier to identify the presence of [$^{13}C_{6}$] lysine in an animal assigned the [$^{2}H_{8}$] valine diet, because of the peak seen at + 6 Da which would not be seen if the animal had consumed only the [$^{2}H_{8}$] valine diet. If both labels are present in a mouse assigned the [$^{13}C_{6}$] lysine diet, the two labelled peptides would overlap completely. Figure 4.8 shows how this overlap would appear. However, due to the isotopologue profiles observed, both labels could be identified if present in the same sample – the ^{13}C peak (M_1) of the [$^{13}C_{6}$] lysine would be significantly more intense than expected due to the addition of the ^{12}C peak (M_0) of the [$^{2}H_{8}$] valine, as would the $^{13}C_2$ peak (M_2) due to the addition of the ^{13}C peak of [$^{2}H_{8}$] valine. This change in peptide profile would indicate the presence of both labels in the urine of a mouse fed the [$^{13}C_{6}$] lysine diet.

4.2.2.3 Calculation of precursor RIAs

The precursor RIA values were calculated for the two mono-valine containing MUP peptides for the mice fed both labelled diets, as in both cases, valine containing and lysine containing peptides were required to determine the presence (or absence) of either label in all samples. The use of endoprotease Lys-C for proteolysis prior to mass spectrometric analysis resulted in all peptides being lysine-terminated.

For each mouse, precursor RIAs for each MUP peptide were calculated as before, and an average RIA was calculated for each mouse, each day, based on the presence of the 'correct' label. This allowed us to see whether the feeding mechanism was allowing the mice to reach their assigned diet, resulting in the mice reaching a labelling maximum. Then, the presence of any 'wrong' label was searched for, to ensure that the feeding mechanism was working in moving feeding traps away from the 'wrong' mice. Taking into account the care required in identifying the presence of 'wrong' label in MUP peptides, none of the mice in any of the three pairs showed any evidence of the 'wrong' label in their urine samples. All MUP peptides

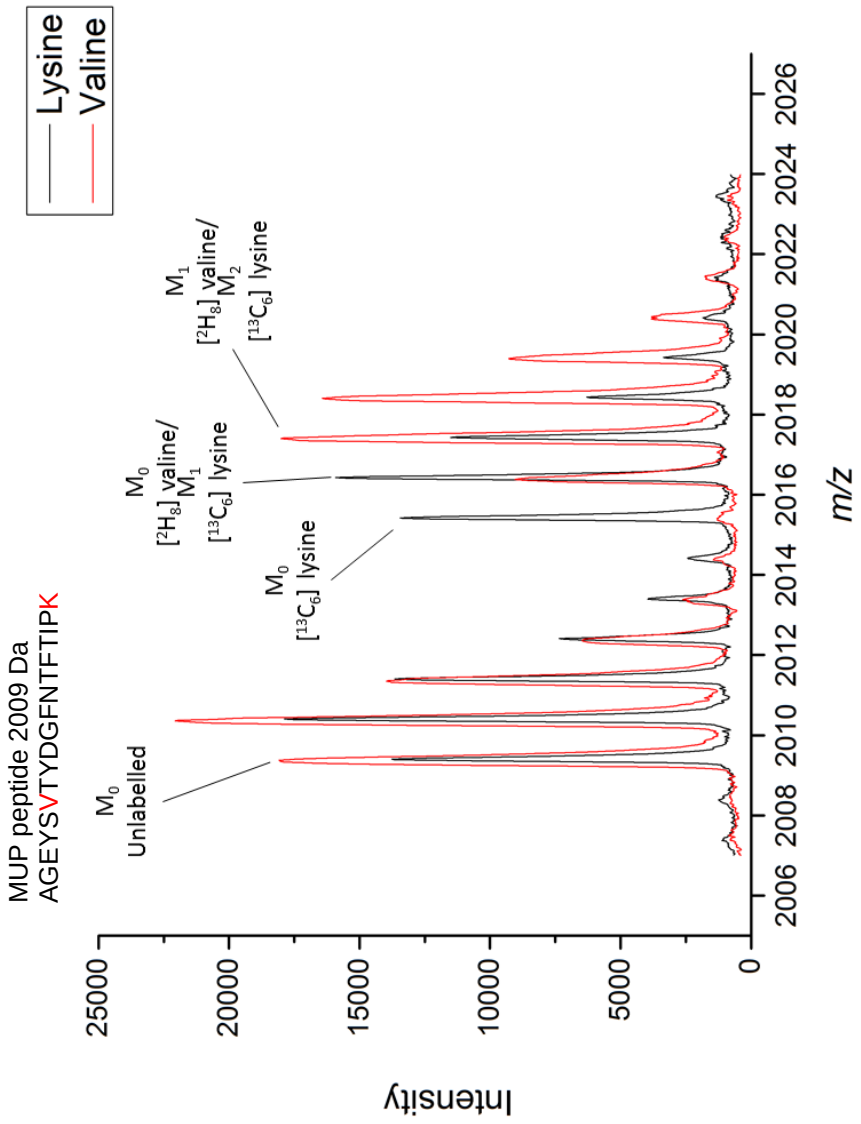


Figure 4.8 The difference between a MUP peptide, containing both a valine residue and a lysine residue, labelled with [$^{2}H_{8}$] valine (red isotopic profile) and with [$^{13}C_{6}$] lysine (black isotopic profile).
The difference in the isotopic profiles seen in MALDI-ToF-MS spectra of urine from BALB/c females fed a [$^{2}H_{8}$] valine diet and a [$^{13}C_{6}$] lysine diet, even though the difference between the two 'heavy' peptide profiles is only 2 Da, is key for ascertaining which label is present in the animal. If the feeding mechanism was not working correctly, both 'heavy' labels would be present in the MUP peptides of a single mouse, and a mixture of the two 'heavy' profiles would be observed in a peptide which contained both a valine and a lysine residue.

observed in the MALDI-ToF-MS spectra were checked to ensure no trace of the [$^{2}H_{8}$] valine peptide in the urine of mice assigned the [$^{13}C_{6}$] lysine diet, and *vice versa*. The absence of any label from the unassigned diet in any mouse confirmed that the feeding mechanism worked successfully in terms of removing the diet from reach when the mouse not assigned to that diet approached.

For the mice assigned the [$^{2}H_{8}$] valine diet, precursor RIA values were calculated, as before, from the unlabelled and [$^{2}H_{8}$] valine labelled peptides. The daily average RIAs for each mouse were averaged again and plotted to determine whether maximum labelling had been reached (Figure 4.9 (a)), with standard deviation calculated for each daily average RIA. Mice appeared to be fully labelled by the time the urine sample was collected on day 1, with variation most likely due to the averaging of RIA values from three different mice.

The same was done for the mice assigned the [$^{13}C_{6}$] lysine diet, but from the unlabelled and [$^{13}C_{6}$] lysine labelled peptides (Figure 4.9 (b)). Again, mice appeared to be fully labelled by day 1 of being fed the labelled diet, but the variation in calculated RIAs was much greater at day 1. This is likely to be due to larger differences in RIA in the three different mice, which could have arisen due to greater differences in the amount of labelled diet eaten by the mice at this time point.

It was concluded that the feeding mechanism designed at Leahurst was suitable in providing pairs of mice differently labelled diets whilst allowing them to be housed in the same cage, which means that pairs of female mice will be able to live together, with unrestricted access to their nestbox and pup litters, and consume differently labelled diets. This means a single label from each female will be able to be passed on to the pups in the communal nest, meaning that maternal investment of each female in the pups should be able to be tracked via the presence of labels in the pups' stomach contents (containing labelled milk passed on from the mother).

4.2.3 Milk labelling pilot

4.2.3.1 Experiment set-up

The kinetics of the incorporation of labelled amino acids via a semi-synthetic diet were established, and a communal nursing cage set-up with a feeding mechanism to enable the feeding of differently labelled diets to two mice has been developed, tested and proved successful. The next aim was to determine whether female mice, housed communally and fed labelled diets, pass the label incorporated into them onto their

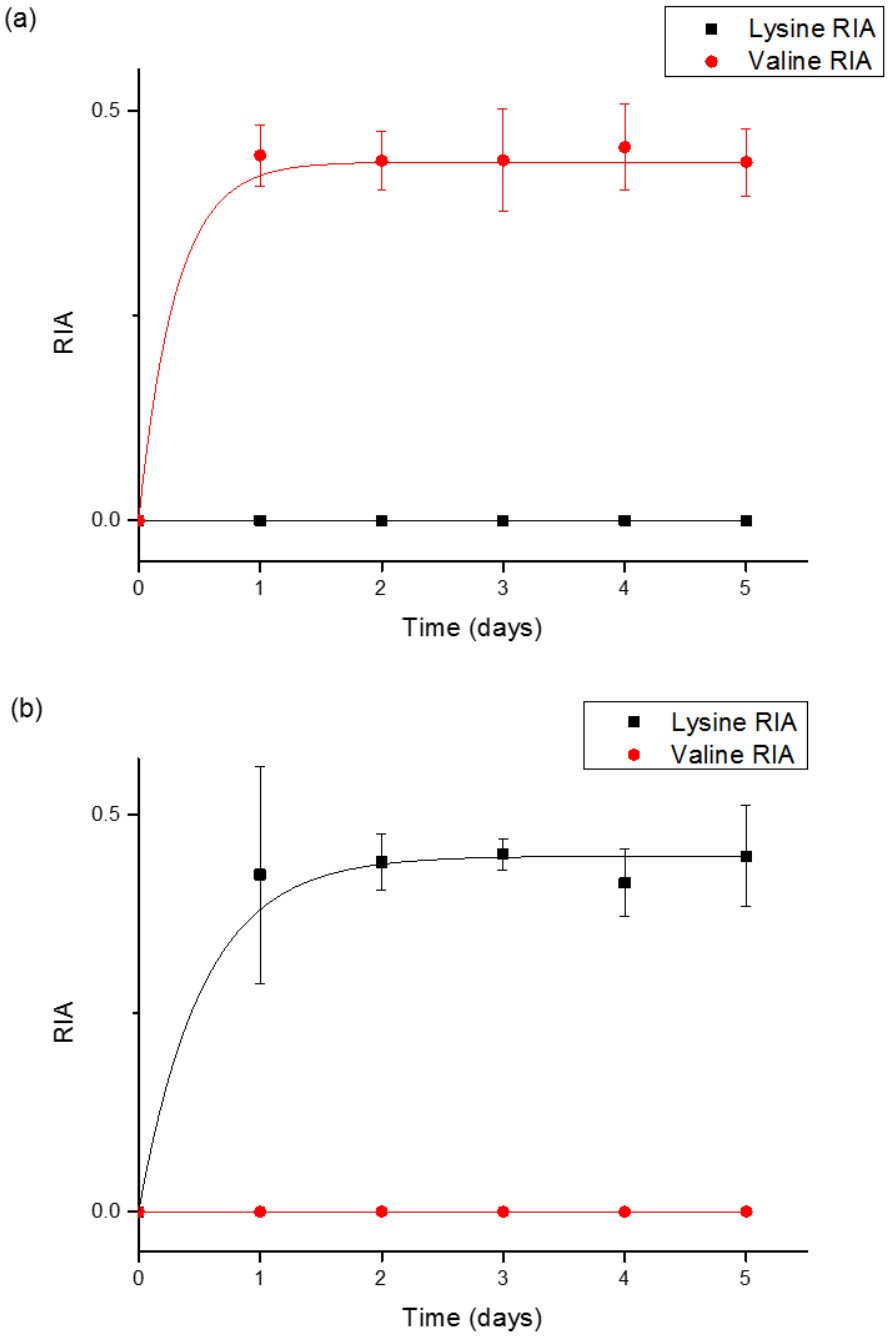


Figure 4.9 The rate of heavy amino acid incorporation over 6 days in (a) three mice fed [$^{2}H_{8}$] valine labelled diet and (b) three mice fed [$^{13}C_{6}$] lysine labelled diet.
Three nests with the feeding mechanism contained two mice in each, each mouse in the pair was assigned either the [$^{2}H_{8}$] valine diet or the [$^{13}C_{6}$] lysine diet. Urine samples were taken from the six BALB/c females every day over the six-day period, subjected to in-solution digestion with Lys-C and analysed using MALDI-ToF-MS. An average daily RIA was calculated from the two most abundant valine and lysine-containing MUP peptides for each mouse, each day, with error bars showing the standard deviation calculated for each daily RIA.

litters of pups via the milk they invest, and the rate at which the label is incorporated into the milk.

A total of six trios (two female BALB/c and one male BALB/c mice in each trio) were set up for breeding at Leahurst by Dr. J. P. Green. The female mice in each trio were previously cage-mates, and were fed an unlabelled diet. After mating, the six pairs of female mice were transferred to the communal nursing cage, continuing with the unlabelled diet. Out of these, three pairs of female mice successfully reproduced and gave birth to litters.

On day 0, when both litters to a pair of females had been born, the female mice then had their diet swapped to a [$^{13}C_6$] lysine labelled diet. Unlike previous experiments, both females were fed the same labelled diet, as the purpose of this experiment was simply to identify the rate at which label is incorporated into the milk of the lactating mothers. Immediately before the female mice were given the labelled diet, a single pup was taken to sample stomach contents – this was to demonstrate that prior to being fed the labelled diet, the mothers' milk contained no label. From this point, all three female pairs were fed the [$^{13}C_6$] lysine labelled diet until day 6 (Figure 4.10).

At time intervals during the course of the experiment, urine was sampled from each mother to ensure the rapid and complete labelling from the diet. Over days 1 - 6, one pup was removed from two of the three communal litters for analysis of stomach contents. From each of these pups, urine samples were taken from their bladder whenever possible, and liver and muscle samples were also taken for future analysis of the kinetics of pup tissue labelling over the course of the six days. Pups were taken from a litter no more than two days on the run in order to minimise disruption in the communal nests. Removing so many pups from a single litter may cause mothers to invest significantly more in their remaining pups (that would be sampled at later time points) than they did in the pups removed from the litter earlier on in the experiment, making meaningful determination of the kinetics of milk labelling challenging. By taking a pup from a communal litter on no more than two consecutive days, the rate of milk labelling can be determined more accurately.

4.2.3.2 Analysis of mothers' urine, pup stomach contents and pup urine

The proteins in the mothers' urine, the milk from the pups' stomach contents and the urine from pups' bladders were digested in-solution with Lys-C and analysed using MALDI-ToF-MS. The digested pup stomach contents proteins and pup urine proteins

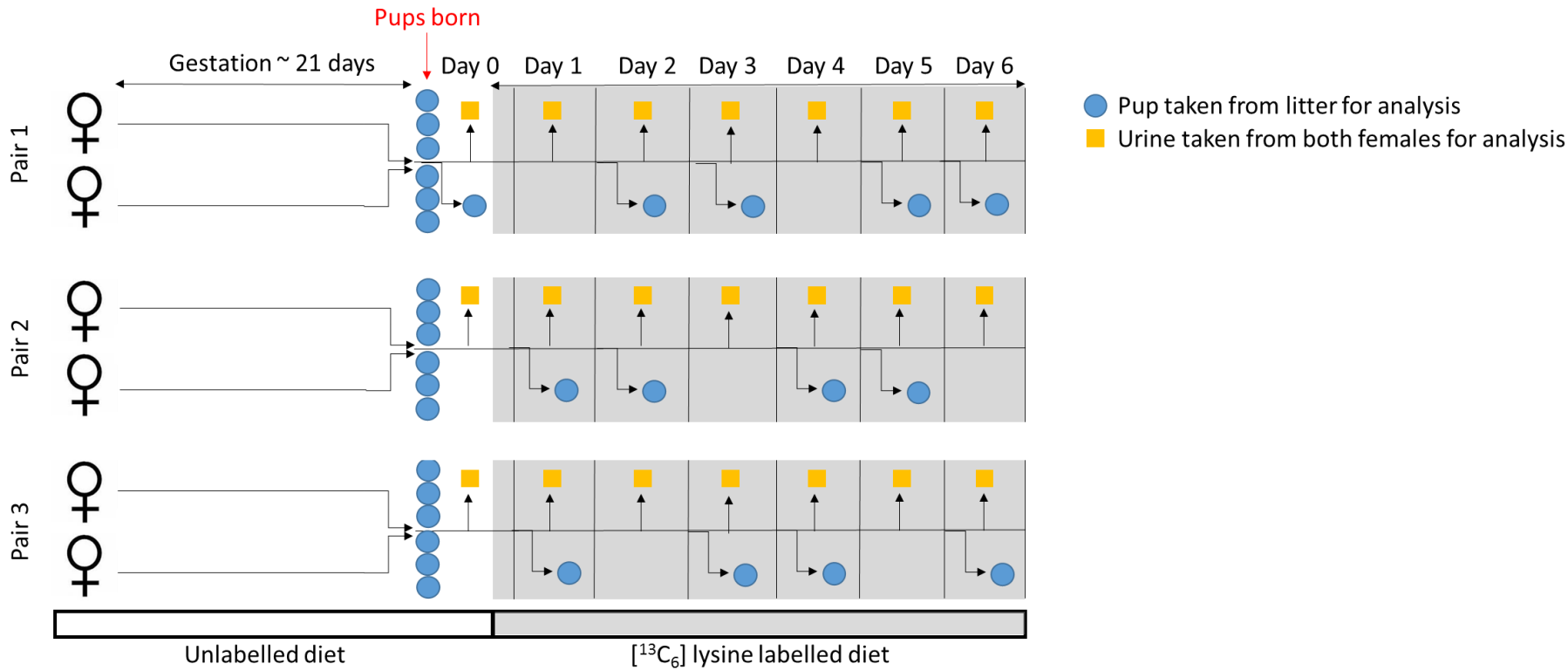


Figure 4.10 The experimental set up at Leahurst for the milk labelling pilot.
Three pairs of pregnant females lived and gave birth in in a communal nursing cage, consuming an unlabelled diet. When all females had given birth, a single pup (blue circle) was taken from one of the communal litters for analysis. All females had their urine sampled (yellow square) at this point. All females were then switched to a [$^{13}C_6$] lysine labelled diet (the part of the experiment where females consumed labelled diet is shaded in grey). Over days 1 – 6, urine from all females was sampled daily for analysis (yellow square). Two pups were taken from one of the three communal litters (blue circle) for analysis each day over the course of the labelling experiment.

were also subjected to LC-MS analysis using a Waters Synapt G2 for protein discovery. LC-MS was chosen for protein discovery over MALDI-ToF-MS because the additional LC step allows the separation of peptides prior to MS analysis, and the use of ESI with a Q-ToF mass spectrometer for MS analysis results in multiply charged peptides which are easier to fragment using CID, and so provide better product ion information for sequence analysis, therefore allowing more accurate peptide identification. For the urine samples from each mother, unlabelled MUP peptides were identified in the mass spectra, along with the corresponding [$^{13}C_6$] lysine labels at a +6 Da shift, and precursor RIAs for the 2471 Da MUP peptide in each sample were calculated using the formula outlined previously. Only one MUP peptide was needed for the calculations as the previous analyses confirmed no variation in the rate of labelling of different MUP peptides.

To identify the peptides (and the proteins from which they were generated) present in the MALDI-ToF mass spectra for milk and pup urine samples, the mass spectra generated by LC-MS were processed and searched against the UniProt *Mus musculus* database using PLGS (v.2.5.2). The top 20 proteins identified in the milk and urine samples are listed in Tables 4.2 and 4.3.

Figure 4.11 shows the mass spectrum of a milk sample from a pup that had been taken just one day after the mothers had begun consuming their [$^{13}C_6$] lysine labelled diet. This mass spectrum is a typical representation of the peptides present in all milk samples. The figure includes a list of the peptides identified, and the proteins from which they were generated, confirmed by cross referencing the peptide masses with those observed in LC-MS analysis. Proteins include LACE1_MOUSE (lactation elevated protein) which is highly expressed in the lactating breast of the female, and in this case has been passed from the mothers to the pups via milk investment. Also present were proteins found specifically in milk, such as alpha-S2-casein-like A and lactadherin. As the lactation elevated protein peptide 2433 Da was consistently the most intense throughout all milk MALDI-ToF-MS spectra, and is specifically expressed in lactating breasts of female mice, this was the peptide used to calculate the rate of labelling in milk.

Figure 4.12 shows the mass spectrum of a urine sample from a pup that had been taken one day after the mothers had begun consuming their [$^{13}C_6$] lysine labelled diet. This mass spectrum is a typical representation of the peptides present in all urine samples. The figure includes a list of the peptides identified, and the proteins from which they were generated, confirmed by cross referencing the peptide masses with

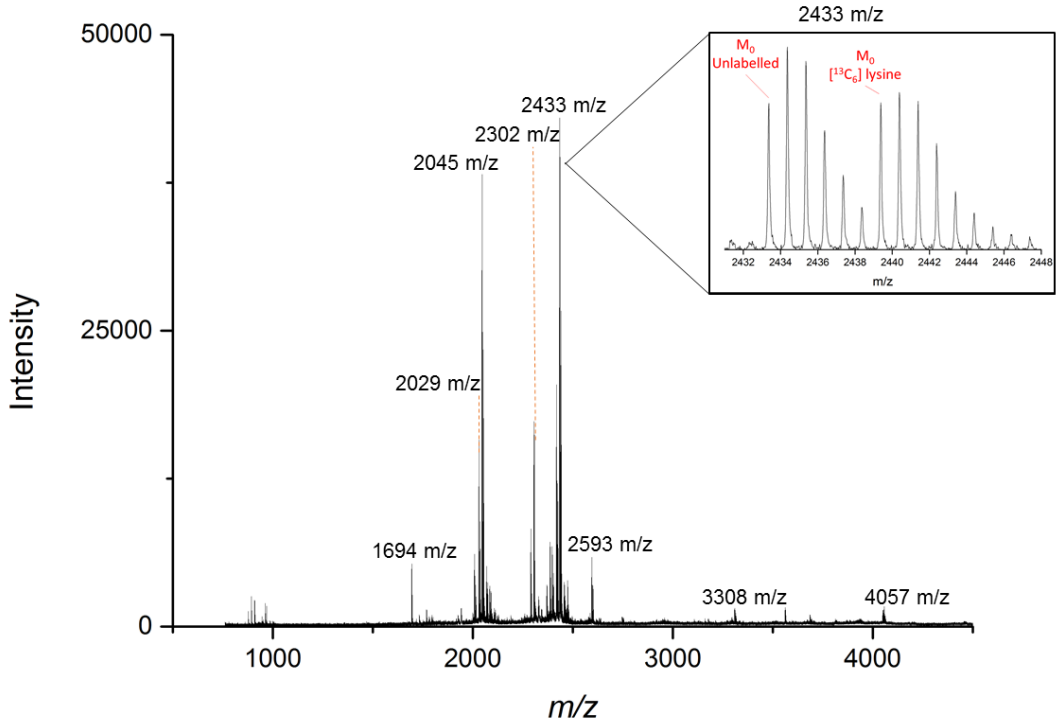
Table 4.2 The top 20 proteins identified in pup stomach contents (milk) samples recovered days 1 – 6 of the experiment.
Raw mass spectra from LC-MS analysis were processed and searched against a database of *Mus musculus* proteins obtained from UniProt, using PLGS.

Protein accession number	Protein name	Score	Sequence coverage (%)	No. of peptides
Q02596	Glycosylation-dependent cell adhesion molecule 1	34802	84.1	8
A0A0G2JGT8	Alpha-S1-casein (Fragment)	33936	100	9
Q02862	Alpha-S2-casein-like A	28263	59.8	12
F8WIP8	Osteopontin	25465	57.3	16
Q921I1	Serotransferrin	18959	38.2	21
E9Q035	Protein Gm20425	18934	21.9	20
P01942	Haemoglobin subunit alpha	18044	71.8	8
P33622	Apolipoprotein C-III	17718	96.0	4
P02664	Alpha-S2-casein-like B	16748	65.0	22
B1ARV3	L-amino acid oxidase 1	14894	44.4	22
Q91VB8	Alpha globin 1	12887	51.4	7
P02088	Haemoglobin subunit beta-1	12694	55.8	4
P07724	Serum albumin	11420	43.4	19
Q00623	Apolipoprotein A-I	11154	61.0	15
P07759	Serine protease inhibitor A3K	10996	59.8	15
Q6JHY2	Submandibular gland protein C	9944	56.2	19
P10598	Beta-casein	9872	55.4	5
P63260	Actin_ cytoplasmic 2	7699	76.0	15
P21956	Lactadherin	5555	29.8	17
P17182	Alpha-enolase	5499	70.3	20

.

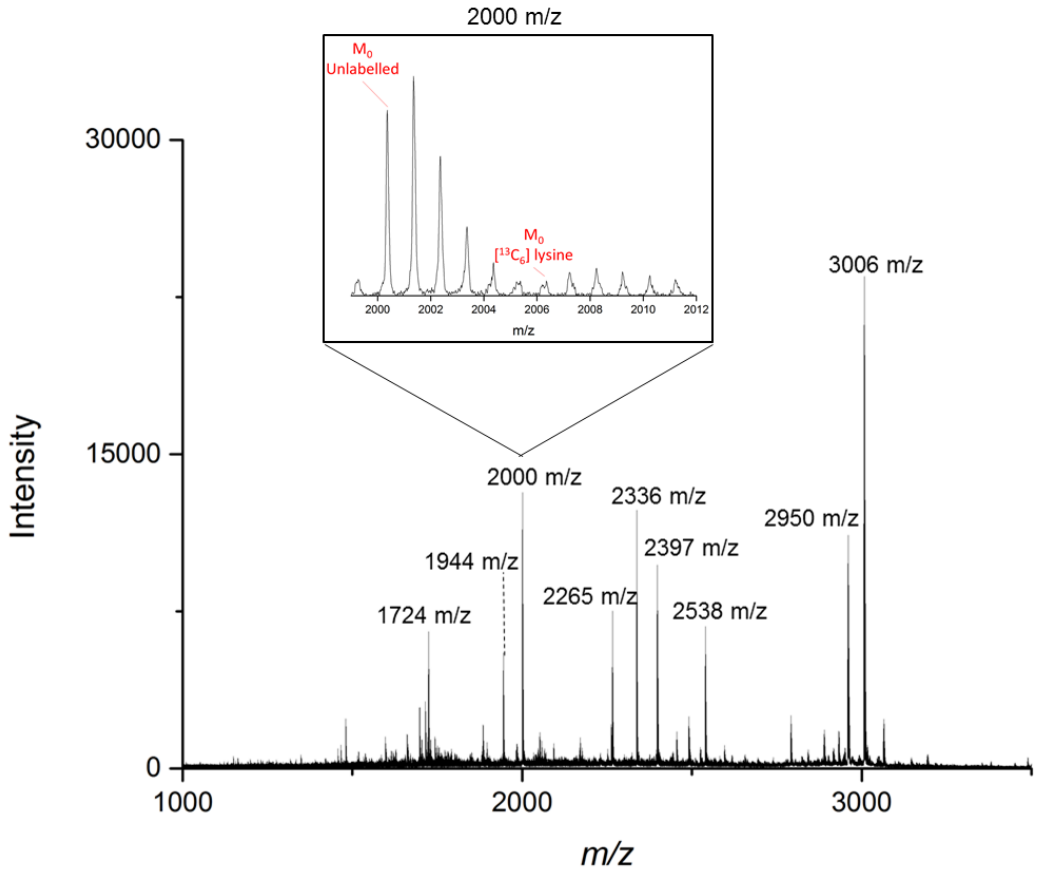
Table 4.3 The top 20 proteins identified in pup urine samples recovered days 1 – 6 of the experiment.
Raw mass spectra from LC-MS analysis were processed and searched against a database of *Mus musculus* proteins obtained from UniProt, using PLGS.

Protein accession number	Protein name	Score	Sequence coverage (%)	No. of peptides
Q7M748	Whey acidic protein	13692	73.3	12
P07724	Serum albumin	3024	63.0	26
Q9DAU7	WAP four-disulfide core domain protein 2	2984	40.4	6
Q02596	Glycosylation-dependent cell adhesion molecule 1	2853	29.8	4
Q9D3N5	Protein 5430402E10Rik	2739	20.8	3
Q9D3H2	MCG117626	2498	34.4	4
P35459	Lymphocyte antigen 6D	2305	22.1	3
B7ZNJ1	Fibronectin	1710	13.3	12
P15947	Kallikrein-1	1657	64.0	9
P02772	Alpha-fetoprotein	1322	39.2	15
A2BHD2	Protein Gm14743	1032	33.3	6
A2AEP0	Protein Obp1b	963	27.5	5
E9Q4P0	KxDL motif-containing protein 1 (Fragment)	805	25.9	7
P29752	Alpha-lactalbumin	740	21.6	2
Q9JKP7	DNA polymerase epsilon subunit 3	720	7.6	2
Q68FH0	Plakophilin-4	620	9.9	6
Q8BHC0	Lymphatic vessel endothelial hyaluronic acid receptor 1	609	4.1	2
P62984	Ubiquitin-60S ribosomal protein L40	605	13.3	3
E9Q5F6	Polyubiquitin-C (Fragment)	570	8.0	2
O09114	Prostaglandin-H2 D-isomerase	567	9.0	2



Peptide (m/z)	Protein Accession	Protein Name	Peptide Sequence
1694	Q02862	Alpha-S2-casein like A	DVTYFPNAHYTRFY(-)
2029	P06728	Apolipoprotein A-IV	ENVDNLHTSMMPLATNLK
2045	P06728	Apolipoprotein A-IV	ENVDNLHTSMM(Ox)PLATNLK
2304	P21956	Lactadherin	MRVSGVMTQGASRAGRAEYLK
2433	Q3V384	Lactation elevated protein 1	RVHFHGFMLDVHRRIHHLK
2593	Q6JHY2	Submandibular gland protein	SDSGSHNLSSGSGSRSNVSTGGEPSDK
3308	P60710	Actin, cytoplasmic 1	AGFAGDDAPRAVEFPSIVGRPRHQGVGMGQK
4057	P33622	Apolipoprotein C-III	TVQDALSSVQESDIAVVARGWMDNHFRFLKGYWSK

Figure 4.11 Peptide mass fingerprint of pup milk proteins taken on the first day of its mother being fed a labelled diet.
Pup milk was homogenised and digested using the homogenisation and in-solution digestion protocols, with Lys-C as the protease. 1µl of digested sample was mixed 1:1 with matrix solution before being spotted onto the MALDI target plate. Mass spectrum was acquired using the method described in Chapter 2. Labelled are the most abundant peptides identified by matching the masses to those identified in the LC-MS discovery analysis of the pup stomach contents (Table 4.2). Inset is an enlargement of the unlabelled and [$^{13}C_6$] lysine labelled lactation elevated protein 1 peptide, at 2433 m/z.



Peptide (m/z)	Protein Accession	Protein Name	Peptide Sequence
1724	P15947	Kallikrien - 1	NSQPWQVAVYRFTK
1944	P07724	Serum albumin	SEIAHRYNDLGEQHFK
2000	P07724	Serum albumin	LGEYGFQNAILVRYTQK
2265	P07724	Serum albumin	APQVSTPTLVEAARNLGRVGTK
2336	P07724	Serum albumin	AWAVARLSQTFPNADFAEITK
2397	P07724	Serum albumin	LCAIPNLRENYGELADCCTK
2538	P07724	Serum albumin	DDNPSLPPFERPEAEAMCTSFK
2950	P07724	Serum albumin	CCAEANPPACYGTVLAEFQPLVEEPK
3006	Unidentified	Unidentified	Unidentified

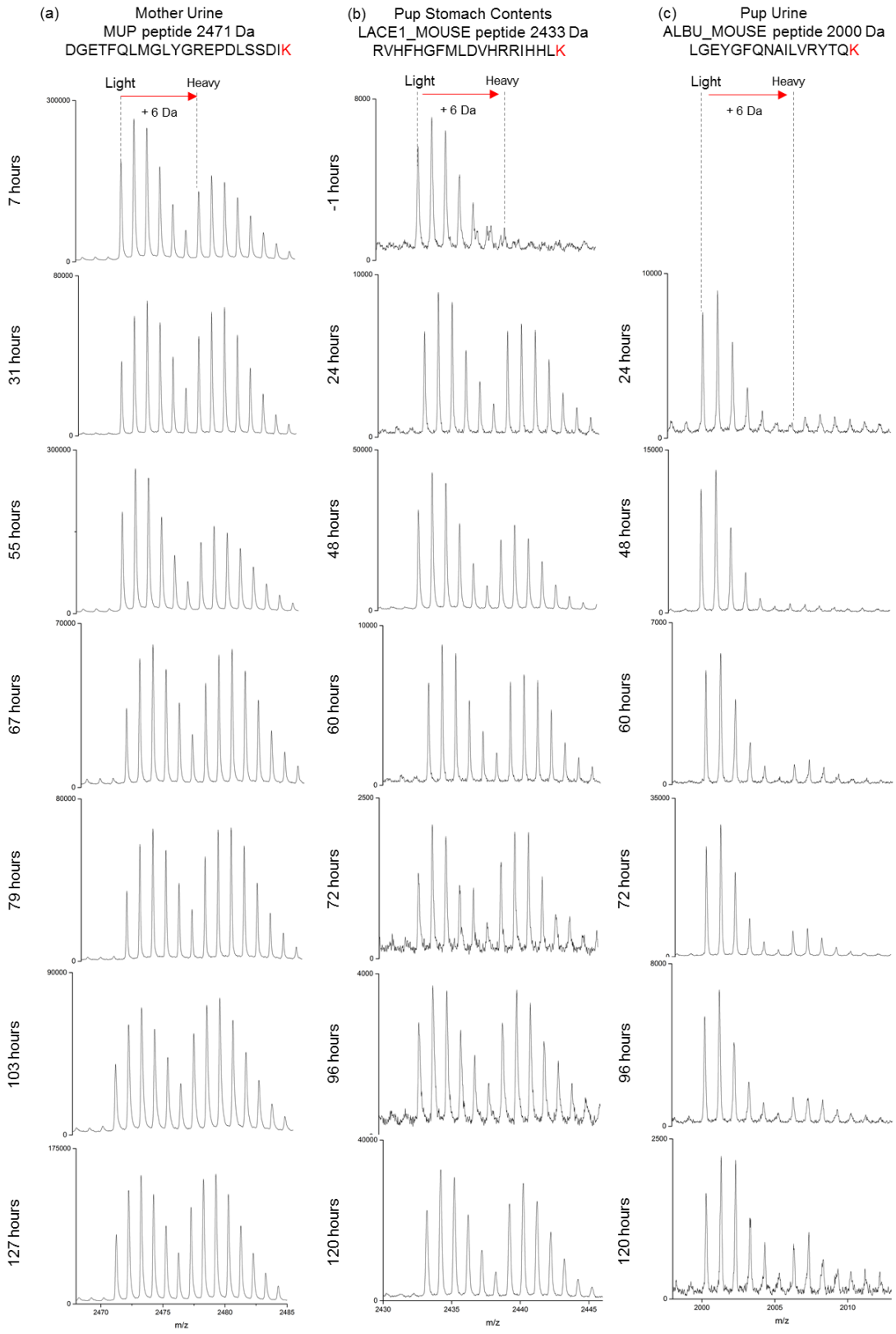
Figure 4.12 Peptide mass fingerprint of pup urine proteins taken on the first day of its mother being fed a labelled diet.
Pup urine was digested using the in-solution digestion protocol, with Lys-C as the protease. 1µl of digested sample was mixed 1:1 with matrix solution before being spotted onto the MALDI target plate. Mass spectrum was acquired using the method described in Chapter 2. Labelled are the most abundant peptides identified by matching the masses to those identified in the LC-MS discovery analysis of the pup urine (Table 4.3). Inset is an enlargement of the unlabelled and [$^{13}C_6$] lysine labelled serum albumin peptide, at 2000 m/z.

those observed in LC-MS analysis. Proteins include ALBU_MOUSE (serum albumin). As all pup urine mass spectra showed intense serum albumin peptides, the serum albumin peptide 2000 Da was used to calculate precursor RIA and therefore the rate of labelling of this protein in pups.

Initially, the unlabelled and [$^{13}C_6$] lysine labelled peptide profiles for MUP peptide 2471 Da (which is the same peptide used to assess the kinetics of [$^{13}C_6$] lysine incorporation into female mice in the first experiment) were observed at different time points over the six days of labelling. Figure 4.13 (a) shows that in the first urine sample taken from a female (mother), taken 7 hours after being provided the [$^{13}C_6$] lysine diet, this MUP peptide was already partially labelled. On the second day of labelling (at 31 hours after being fed the labelled diet), this MUP peptide appeared fully labelled, and remained fully labelled for the remainder of the experimental period, apart from a small dip in labelling on day 2. This is likely to be due to variation in the amount of diet consumed. Apart from this, the rate of labelling was in agreement with the rate seen in the previous proof-of-principle experiment, where the mice fed the [$^{13}C_6$] lysine diet appeared to have a fully labelled precursor pool after one day of being fed the labelled diet. For the calculation of precursor RIAs in MUPs, and therefore calculation of the labelling of the precursor pool in the liver, an average RIA of two MUP peptides in each sample taken from each of the six females was calculated.

The unlabelled and [$^{13}C_6$] lysine labelled peptide profiles for the lactation elevated protein peptide 2433 Da in pup stomach contents (milk) were observed at different time points over the six days of the experiment. All milk samples were recovered at the same time every day, except for the very first sample – this was from the pup taken from the communal litter immediately before the mothers were provided with the [$^{13}C_6$] lysine diet, in order to prove that no label was present in the mothers' milk prior to mice being fed the labelled diet. In Figure 4.13 (b), the mass spectrum of the 2433 Da LACE1_MOUSE peptide, from the milk sample taken 1 hour before labelling, shows the presence of only the unlabelled peptide, with no labelled peptide at the characteristic +6 Da shift. All milk samples taken after this point were from pups whose mothers were being fed the labelled diet. Observation of the mass spectra of the LACE1_MOUSE 2433 Da peptide over the course of the experiment indicates a similar rate of labelling to the labelling of MUPs seen in the mothers' urine. On the first day of mothers being fed the [$^{13}C_6$] lysine diet, the milk peptide found in the pups' stomach contents appeared fully labelled, with the +6 Da labelled peptide having a

Figure 4.13 Incorporation of heavy labelled amino acids over the course of 6 days, shown by the relative abundances of heavy/light peptide profiles in urine from a BALB/c female (a), stomach contents of a pup (b) and urine from a pup (c), where the female (mother of the pups) was fed a diet labelled with [$^{13}C_6$] lysine.
Urine samples were taken from the females daily over a six-day period. Shown are the unlabelled ('light') and labelled ('heavy') MUP peptide profiles taken from urine collected at intervals after the labelled diet being introduced. The 'heavy' lysine MUP peptide is 6 Da heavier than the light MUP peptide, due to the incorporation of six ^{13}C into the lysine residue (a). The stomach contents and urine samples were taken from a different pup at each time of sampling, shown are the 'light' and 'heavy' profiles from a lactation elevation protein peptide (b) and from a serum albumin peptide (c).


(a)
Mother Urine
MUP peptide 2471 Da
DGETFQLMGLYGREPDLSSDIK
(b)
Pup Stomach Contents
LACE1_MOUSE peptide 2433 Da
RVHFHGFMLDVHRRIHHLK
(c)
Pup Urine
ALBU_MOUSE peptide 2000 Da
LGEYGFQNAILVRYTQK
Light
Heavy
+ 6 Da
7 hours
31 hours
55 hours
67 hours
79 hours
103 hours
127 hours
-1 hours
24 hours
48 hours
60 hours
72 hours
96 hours
120 hours
m/z

similar intensity to the unlabelled peptide. This suggested that the proteins in the mothers' milk were labelling at the same rate as MUPs in their urine. However, the proteins in the milk samples recovered on days 2 and 3 of the experiment appeared only partially labelled – this could be due to the fact that pups are taken from different litters on each day of the experiment, so there may be variation in the investment of mothers in different communal litters, or variation in the amount of labelled diet consumed by the different mothers, and therefore variation in the amount of label passed onto the pups. Another thing that may have caused this dip in milk labelling is the time of day at which pup samples were taken – all pup sampling took place at 9 am of each day, which is at the beginning of the active period for the mice – mothers will have been inactive during their 'day' period overnight, and will have been consuming less diet over this time and investing less in their litters, causing the 'dip' in labelling which was also seen in the morning urine samples taken in the initial kinetics experiment. In this experiment, the mothers' urine samples were collected later in the day, between 4 – 5pm. This is after the 'night' period, where mice are most active and so will be eating more of the labelled diet, which is why mothers' urine samples appeared to have reached a maximum labelling plateau. The differences in the time points where mothers' urine and mothers' milk in pups were sampled is a likely explanation for the 'dip' of labelling seen in the pups' stomach contents on days 2 and 3, along with variation in diet consumed by the mother (as a dip in the labelling of mothers' urine was also seen on day 2). Taking this into account, and referring back to the initial kinetics experiment and the variation of labelling seen at different times during the course of a day, it seems as though the proteins in the mothers' milk seen in the pups' stomach reaches a similar labelling maximum at a similar rate to that seen in the mothers' urine. The calculation of the average precursor RIA values for this LACE1_MOUSE protein peptide in milk for each sampled pup will enable us to determine the rate of incorporation of [$^{13}C_6$] lysine from diet into the mothers' milk, the milk (and therefore label) which is then passed onto pups.

Figure 4.13 (c) shows the unlabelled and labelled serum albumin (2000 Da) peptide profiles observed in the MALDI-ToF-MS analysis of the urine recovered from the pups' bladders. No urine sample was recovered from the pup taken immediately before the mothers were given the [$^{13}C_6$] lysine diet, so there are only analyses for the pup urine samples on days 1 – 6 of the experiment. From the MALDI-ToF-MS analysis, the 2000 Da peptide showed much slower rates of label incorporation than the MUP peptide in mother's urine and the LACE1_MOUSE peptide in milk. At no point in the experiment does the ALBU_MOUSE peptide in pup urine appear fully

labelled. The incorporation of the label into the pup urine, although very slow, increases over the course of the six days that the pups' mothers are being fed the labelled diet. In order to assess the rate of label incorporation into the pups due to their mothers' investment, average precursor RIA values were calculated for the ALBU_MOUSE 2000 Da peptide and the ACSM3_MOUSE in each pup sample taken over the course of the six day experiment.

The average precursor RIAs were calculated as previously described for the mothers' urine samples, the pup's stomach contents and the pup's urine samples at each time point and are shown in Figure 4.14. Figure 4.14 (a) shows that the precursor pool in the mothers' liver has reached a labelling maximum on day 1 of being fed the [$^{13}C_6$] lysine labelled diet (31 hours after the diet had been introduced). The calculated RIA is slightly below the maximum on day 2 of labelling (55 hours after the diet had been introduced), before returning to the plateau for the remaining days of the experiment. Referring back to Figure 4.13, the 'dip' in labelling is also evident in the mass spectrum of the MUP peptide in the urine taken at this time. The variation in the RIA values calculated in the samples from the different mothers suggests that at this point, the differences in the amount of diet consumed by each mouse is the likely cause of this 'dip' in labelling.

Milk proteins in the pup's stomach contents appeared to have labelled at a similar rate and to a similar maximum to the mothers' urine samples, with relatively little variation in the RIA values calculated from different proteins in different pups. Figure 4.14 (b) shows that label incorporation appears slightly less in samples taken on day 3, which is likely to be due to the dip in labelling seen in mothers on day 2 due to variation in diet consumption. The labelling of milk proteins appears to reach the same maximum as mothers' urine samples from day 4 onwards, suggesting that the full labelling of the milk precursor pool takes slightly longer than the labelling of the liver precursor pool in the mothers. As the milk samples were taken from the pups' stomachs, it can be confirmed that the [$^{13}C_6$] lysine is being successfully passed on from the mothers to their pups.

In order to assess whether the label is then being incorporated into the nursed pups, the precursor RIAs were calculated in proteins from the urine samples obtained from pups. Figure 4.14 (c) shows that the incorporation of label into the pup is much slower, as would be expected, and at no point during the experiment has it appeared to reach a maximum labelling plateau. This is to be expected because the label

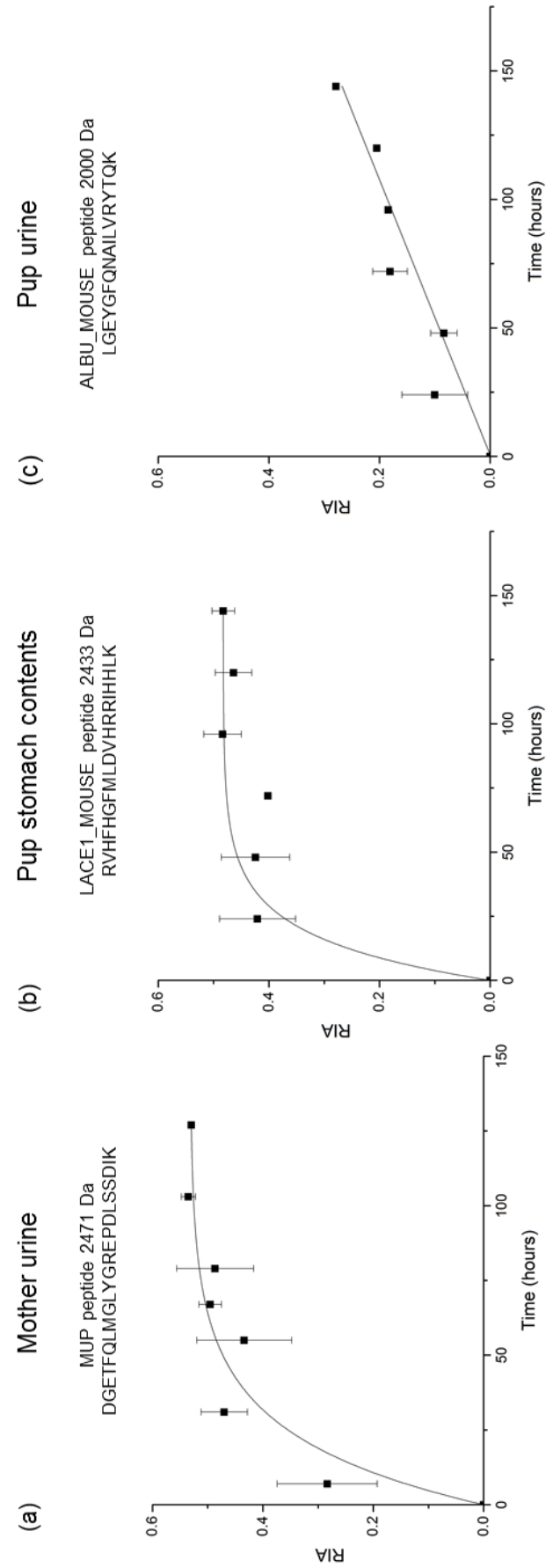


Figure 4.14 The rate of heavy amino acid incorporation over 6 days in (a) urine from BALB/c females, (b) stomach contents of pups and (c) urine from pups, where the female (mother of the pups) was fed a diet labelled with [$^{13}C_6$] lysine.
Urine samples were taken from three BALB/c females at daily over the period, subjected to in-solution digestion with Lys-C and analysed using MALDI-ToF. An average RIA for the three mice calculated from the most abundant lysine-containing peptide for each day, with error bars showing the standard deviation calculated for each daily RIA (n = 6). Pup stomach contents (n = 2) (b) and urine (n = 2 on first three days; n = 1 on final three days) (c) was taken from at least one pup at each interval over the period, and analysis and RIA calculations were carried out in the same way as the mothers' urine (error ± SD).

incorporation into the pups is much slower because of their rate of growth; at this age, protein turnover rates are maximal (Waterlow *et. al*. 1978). The rate of protein synthesis is higher than the rate of protein degradation, and so the label is being diluted by a greater amount of newly synthesised unlabelled protein present in the precursor pool. The calculated precursor RIAs in urine increase throughout the experiment, which indicates that the label is being successfully incorporated into the nursed pups, just at a much slower rate than dietary incorporation into their mothers. At the end of the 6 days, precursor RIAs of the serum albumin peptide have reached a value of around 0.3 – lower than the RIAs reached in proteins in mothers' urine and pups' stomach contents (approximately 0.5), but relatively significantly labelled nonetheless.

4.2.3.3. Analysis of liver and muscle samples from pups

From each pup sampled throughout the 6 day experiment, liver and muscle samples were taken to assess the rate of label incorporation into these two tissues. In general, it would be expected that proteins in the pup liver would label quicker than proteins in muscle, as the muscle is less metabolically active (Claydon *et al.* 2012). It is known that different proteins in a sample will have different rates of label incorporation due to them having different rates of turnover (Waterlow *et al.* 1978). The analysis of RIA in proteins with different turnover rates in pup tissue samples can give information regarding investment from mothers at different time points during communal nursing – for example, a high turnover protein in a pup tissue will degrade and return to the protein precursor pool quickly, meaning that upon introduction of a labelled amino acid to the precursor pool, previously synthesised (unlabelled) proteins degrade quickly and newly synthesised (labelled) proteins take their place. Upon these degrading, labelled amino acids are returned to the precursor pool as well as those being introduced via diet, causing complete labelling of the precursor pool to occur quickly. These proteins therefore give an indication of the recent label investment from mothers. The opposite is true for a low turnover protein, and so these proteins give a picture of the label investment from mothers over a longer time period.

In order to assess the rates of turnover of a number of proteins in pup liver and muscle samples, samples were homogenised, digested with Lys-C and analysed using LC-MS. LC-MS was chosen for protein discovery over MALDI-ToF-MS once again due to increased protein identification accuracy. Raw mass spectra for all pup liver and muscle samples were imported into Waters PLGS (v 2.5.2) software, and peptides (and the proteins from which they were generated) were identified by searching

against a UniProt database of *Mus musculus* proteins. The top 20 protein hits that were identified in all liver mass spectra are outlined in Table 4.4, and the top 20 protein hits that were identified in all muscle mass spectra are outlined in Table 4.5. For each of the liver samples obtained from two pups on the final day (day 6) of the experiment, the RIA values for the three highest scoring peptides for each of the 20 identified proteins were calculated. The average RIA for each protein was calculated from the three peptide RIA values. The same was done for each of the muscle samples obtained from the two pups on the final day of the experiment.

The average RIAs calculated for the top 20 liver proteins from the two pups are shown in Figure 4.15, and those calculated for the top 20 muscle proteins from the same pups are shown in Figure 4.16. The calculated average RIAs for the proteins in the liver samples are similar between the two different pup samples, with the lowest RIA values calculated for the two haemoglobin proteins (subunits beta –1 and beta –2) in both samples, and the highest RIA value seen in the non-specific lipid transfer protein. In most proteins, there was little variation between the RIAs calculated for each peptide, with the exception of the highest RIA non-specific lipid transfer protein, protein disulphide-isomerase and L-lactate dehydrogenase A chain (Figure 4.17). Maximum protein RIAs range widely in the liver samples, from approximately 0.15 in the apparently lowest turnover protein to around 0.47 in the highest turnover protein (which appears to have reached a labelling maximum). In the muscle samples, the calculated average RIAs for each protein were relatively similar between the two different pup samples, with the exception of parvalbumin alpha (0.37 in pup ID 36313, 0.29 in pup ID 36314) (Figures 4.16 And 4.18). As with the liver samples, the lowest calculated RIA was in haemoglobin subunit beta -2, and the highest calculated RIA in the muscle samples was in serotransferrin. Aside from creatine kinase M-type in both muscle samples, and parvalbumin alpha in the first pup sample, there was little variation between the RIAs calculated for each peptide in each protein (Figure 4.18).

The average precursor RIA values were calculated using two different peptides from each protein, and were calculated for glutamate dehydrogenase 1, mitochondrial (a liver protein with a low RIA calculated on the final day of the experiment), non-specific lipid transfer protein (cytoplasmic in the liver, with the highest RIA calculated RIA on the final day), beta-enolase (a muscle protein with the second-lowest RIA calculated on the final day of the experiment) and parvalbumin alpha (a muscle protein with the highest calculated RIA) for each day of the experiment to assess the rate of label incorporation into these proteins and whether they had reached a maximum labelling

Table 4.4 The top 20 common proteins identified in all pup liver samples recovered days 1 – 6 of the experiment.
Raw mass spectra from LC-MS analysis were processed and searched against a database of *Mus musculus* proteins obtained from UniProt, using PLGS.

Protein accession number	Protein name	Score	Sequence coverage (%)	No. of peptides
P02088	Haemoglobin subunit beta-1	12045	76.9	19
P02089	Haemoglobin subunit beta-2	11656	76.9	21
P12710	Fatty acid-binding protein, liver	4044	55.9	8
B1AXW4	Peroxiredoxin-1	3570	71.5	8
P26443	Glutamate dehydrogenase 1, mitochondrial	2521	22.6	10
Q9QXD6	Fructose-1,6-bisphosphatase 1	2451	17.2	6
Q05816	Fatty acid-binding protein, epidermal	1259	38.5	4
P06151	L-lactate dehydrogenase A chain	1224	16.3	5
P09103	Protein disulfide-isomerase	1106	14.7	8
P99027	60S acidic ribosomal protein P2	1011	53.0	4
Q64433	10 kDa heat shock protein, mitochondrial	817	33.3	3
P16460	Argininosuccinate synthase	717	14.6	7
Q8C196	Carbamoyl-phosphate synthase [ammonia], mitochondrial	618	14.4	18
O35490	Betaine--homocysteine S-methyltransferase 1	603	20.1	7
P54869	Hydroxymethylglutaryl-CoA synthase, mitochondrial	563	8.1	5
P07724	Serum albumin	497	13.3	9
B1ATY1	Actin, cytoplasmic 2	493	24.2	3
Q6PHC1	Enolase	468	12.6	3
Q8QZT1	Acetyl-CoA acetyltransferase, mitochondrial	361	9.7	4
P32020	Non-specific lipid-transfer protein	355	13.7	9

Table 4.5 The top 20 common proteins identified in all pup muscle samples recovered days 1 – 6 of the experiment.
Raw mass spectra from LC-MS analysis were processed and searched against a database of *Mus musculus* proteins obtained from UniProt, using PLGS.

Protein accession number	Protein name	Score	Sequence coverage (%)	No. of peptides
P32848	Parvalbumin alpha	67264	82.7	20
P17751	Triosephosphate isomerase	30667	52.8	14
P02089	Hemoglobin subunit beta-2	26660	64.6	12
P07310	Creatine kinase M-type	24088	38.1	17
P21550	Beta-enolase	21418	57.1	22
P05064	Fructose-bisphosphate aldolase A	15801	34.1	11
E9Q5U3	Glyceraldehyde-3-phosphate dehydrogenase	10702	50.8	13
P16065	Carbonic anhydrase 3	8386	39.6	8
Q9R0Y5	Adenylate kinase isoenzyme 1	7234	56.2	9
P68134	Actin, alpha skeletal muscle	6817	40.3	11
P11499	Heat shock protein HSP 90-beta	5682	27.5	14
P09411	Phosphoglycerate kinase 1	4672	49.9	20
P52480	Pyruvate kinase PKM	4272	26.4	12
Q6PHC1	Alpha-enolase	4254	51.9	14
Q921I1	Serotransferrin	3260	11.2	8
P62631	Elongation factor 1-alpha 2	2091	22.2	6
Q5SX49	Profilin	1646	61.6	5
P14152	Malate dehydrogenase	1207	36.2	11
O70250	Phosphoglycerate mutase 2	1202	40.3	8
P31001	Desmin	892	27.5	7

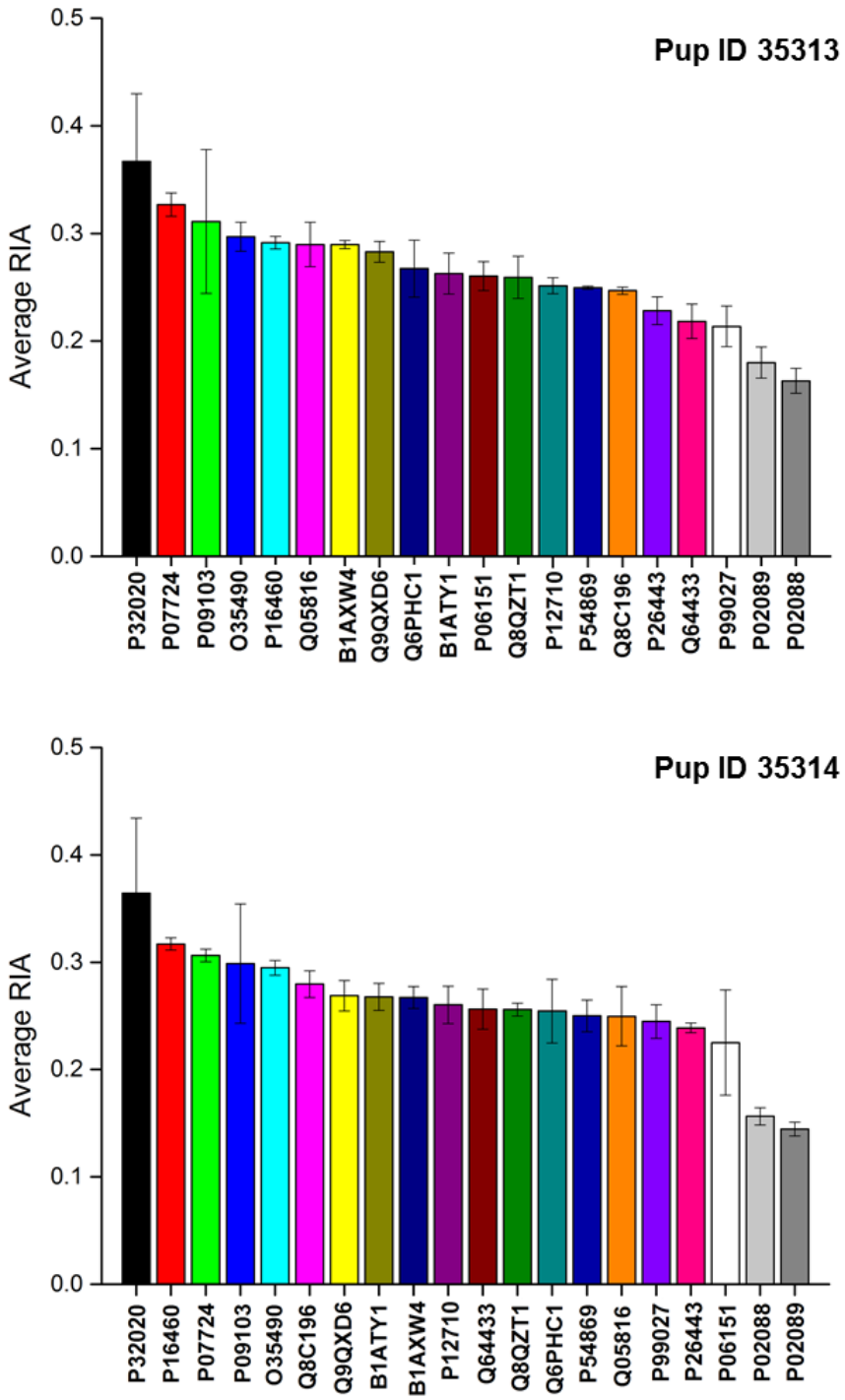


Figure 4.15 The average RIA calculations of the top 20 proteins in pup liver, common in both samples.
The pup liver samples taken from two different pups on the final day of the milk labelling pilot were homogenised and digested with Lys-C as per the methods described in Chapter 2. The digests were analysed using LC-MS and mass spectra were searched for matches in a mouse database using PLGS. For each of the proteins identified here, average RIAs were calculated from the intensities of the top three peptides identified in each protein (error ± SD).

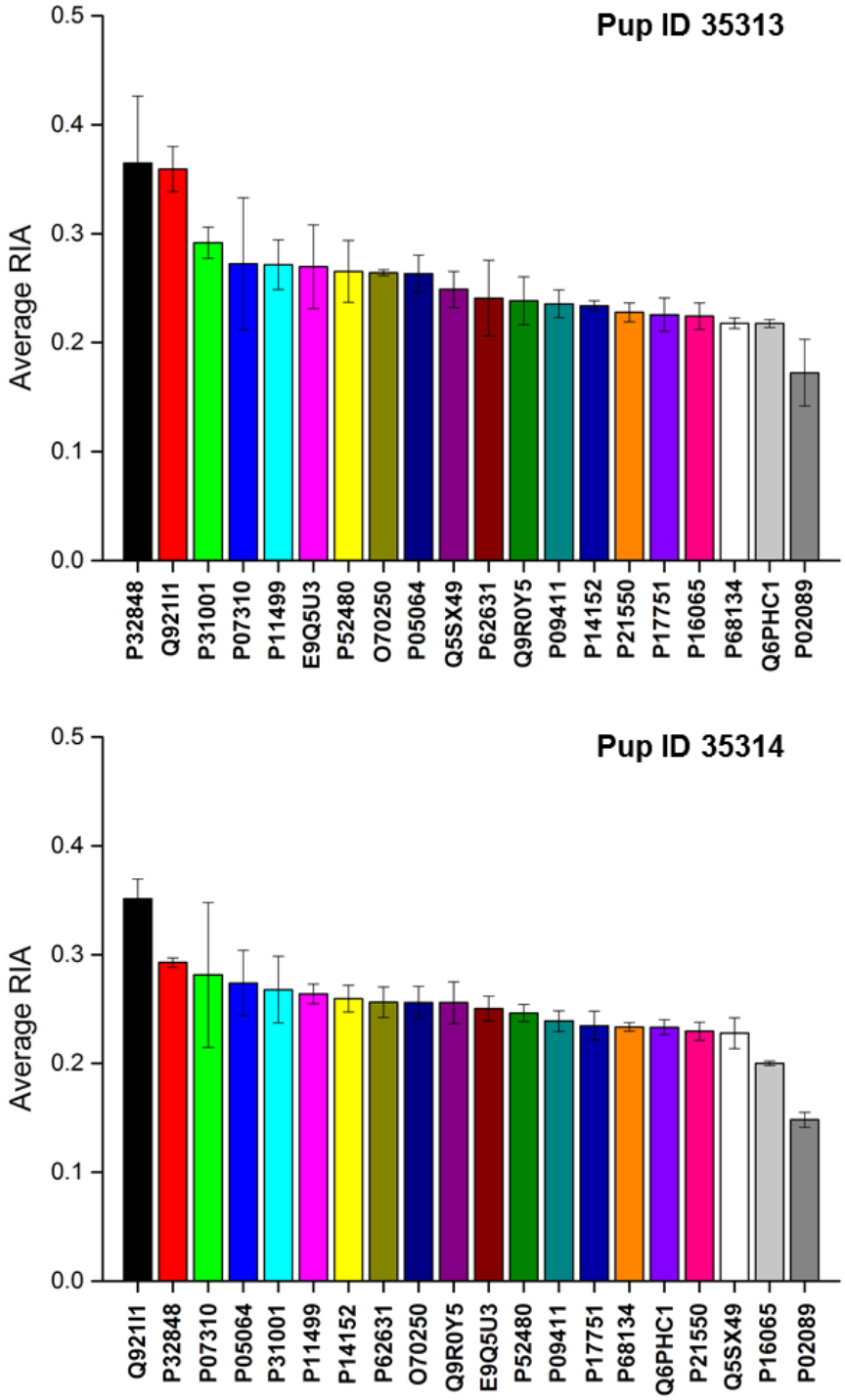


Figure 4.16 The average RIA calculations of the top 20 proteins in pup muscle, common in both samples.
The pup muscle samples taken from two different pups on the final day of the milk labelling pilot were homogenised and digested with Lys-C as per the methods described in Chapter 2. The digests were analysed using LC-MS and mass spectra were searched for matches in a mouse database using PLGS. For each of the proteins identified here, average RIAs were calculated from the intensities of the top three peptides identified in each protein (error ± SD).

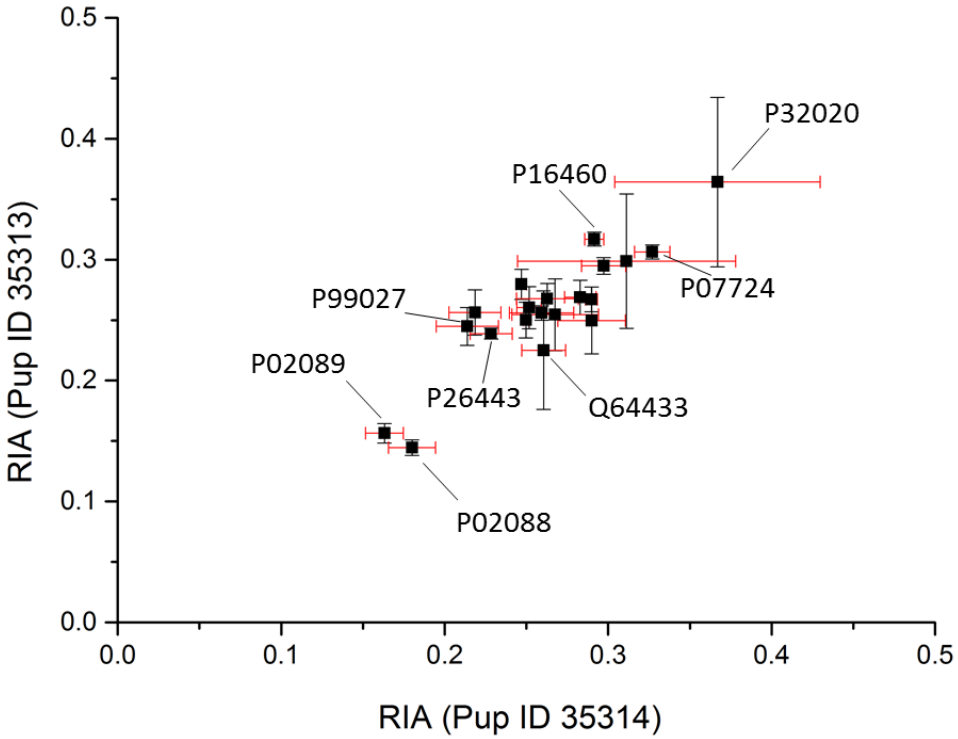


Figure 4.17 The average RIA calculations of the top 20 proteins in pup liver, seen in the both of the pup samples, taken on the final day of the 6-day experiment.
The pup liver samples taken from two different pups on the final day of the milk labelling pilot were homogenised and digested with Lys-C as described in Chapter 2. The digests were analysed using LC-MS and mass spectra were searched for matches in a mouse database using PLGS. For each of the proteins identified here, average RIAs were calculated from the intensities of the top three peptides identified in each protein (error ± SD). Labelled are the proteins with RIA values higher or lower than the majority of the top 20 proteins identified, which may be suitable for analysis for the tracking of investment at different time points in the communal nursing study.

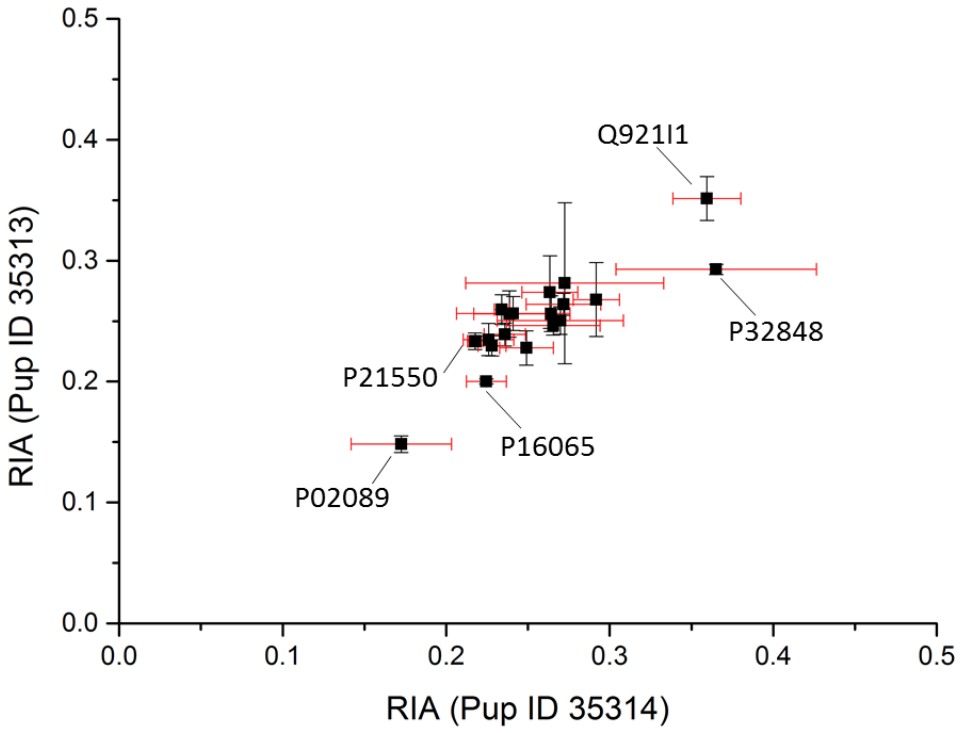


Figure 4.18 The average RIA calculations of the top 20 proteins in pup muscle, seen in the both of the pup samples, taken on the final day of the 6-day experiment.
The pup muscle samples taken from two different pups on the final day of the milk labelling pilot were homogenised and digested with Lys-C as described in Chapter 2. The digests were analysed using LC-MS and mass spectra were searched for matches in a mouse database using PLGS. For each of the proteins identified here, average RIAs were calculated from the intensities of the top three peptides identified in each protein (error ± SD). Labelled are the proteins with RIA values higher or lower than the majority of the top 20 proteins identified, which may be suitable for analysis for the tracking of investment at different time points in the communal nursing study.

plateau at any point during the six days (Figures 4.19 – 4.21). Figure 4.21 shows that each of the two lower turnover proteins label at a similar rate, with the low turnover liver protein (glutamate dehydrogenase 1, mitochondrial in liver) and low turnover muscle protein (beta-enolase) incorporating the label at the slowest rate, reaching a maximum RIA of approximately 0.21 on the final day of the experiment. The higher turnover liver protein (non-specific lipid transfer protein) reached an RIA of 0.36 on the final day of the experiment, and the higher turnover muscle protein (parvalbumin alpha) reached an RIA of 0.29, meaning that these higher turnover proteins were incorporating the label at a faster rate than the lower turnover proteins. All proteins showed an almost linear increase in label incorporation throughout the experiment, with no sign of labelling reaching a plateau at any point over the time period.

From this experiment, it was determined that after giving birth in a communal nest, females were able to consume a [$^{13}C_6$] lysine labelled diet and have that label fully incorporated into them within the space of 2 days, and that this label was successfully passed on to their pups via their milk, which labelled at a similar rate and to a similar extent to their livers. The pups' consumption of the fully labelled milk resulted in the label becoming incorporated into their tissues. Although the rate at which pups were labelled was much slower than their mothers' proteins in these tissues labelled at a significant enough rate to deem this method suitable to track investment in the pups from their mothers.

4.2.4 Tracking investment in pups

Communally housed female mice, who are successfully fed different labelled diets through the development of a suitable feeding mechanism, pass the label incorporated into themselves onto their pups via their milk, and these labels are then incorporated into the pups' tissues. The assessment of labelled amino acid incorporation into pup tissues by calculating protein precursor RIAs can be used as a way to track the investment from the pup's mother. In the milk labelling study, both female mice were fed a diet labelled with [$^{13}C_6$] lysine. The aim of the final communal nursing experiment is to track the investment from each mother in the communal litter, and so each mother is required to consume a differently labelled diet (via the established feeding mechanism (Figure 4.9)). The two different labels will then incorporate into each mother, and will be passed on to the communal litter via their milk, and the labels will be passed onto whichever pups they invest in (Figure 4.22). The hypothesis is that if a pup receives investment from both mothers, both labels

Figure 4.19 Incorporation of [$^{13}C_6$] lysine labelled amino acids over the course of 6 days, shown by the relative abundances of heavy/light peptide profiles in liver samples from pups. (a) the incorporation of [$^{13}C_6$] lysine into the 1583.8 Da peptide of the low turnover protein glutamate dehydrogenase 1, mitochondrial (b) the incorporation of [$^{13}C_6$] lysine into the 916.5 Da peptide of the high turnover protein non-specific lipid transfer protein.
Liver samples were taken from different pups on each day of the six-day experiment. Samples were homogenised, digested in-solution with Lys-C and subjected to LC-MS analysis. Shown are the unlabelled ('light') and labelled ('heavy') peptide profiles of a glutamate dehydrogenase 1, mitochondrial peptide (a) and a non-specific lipid transfer protein peptide (b), confirming the incorporation of [$^{13}C_6$] lysine into these proteins over the 6-day period.

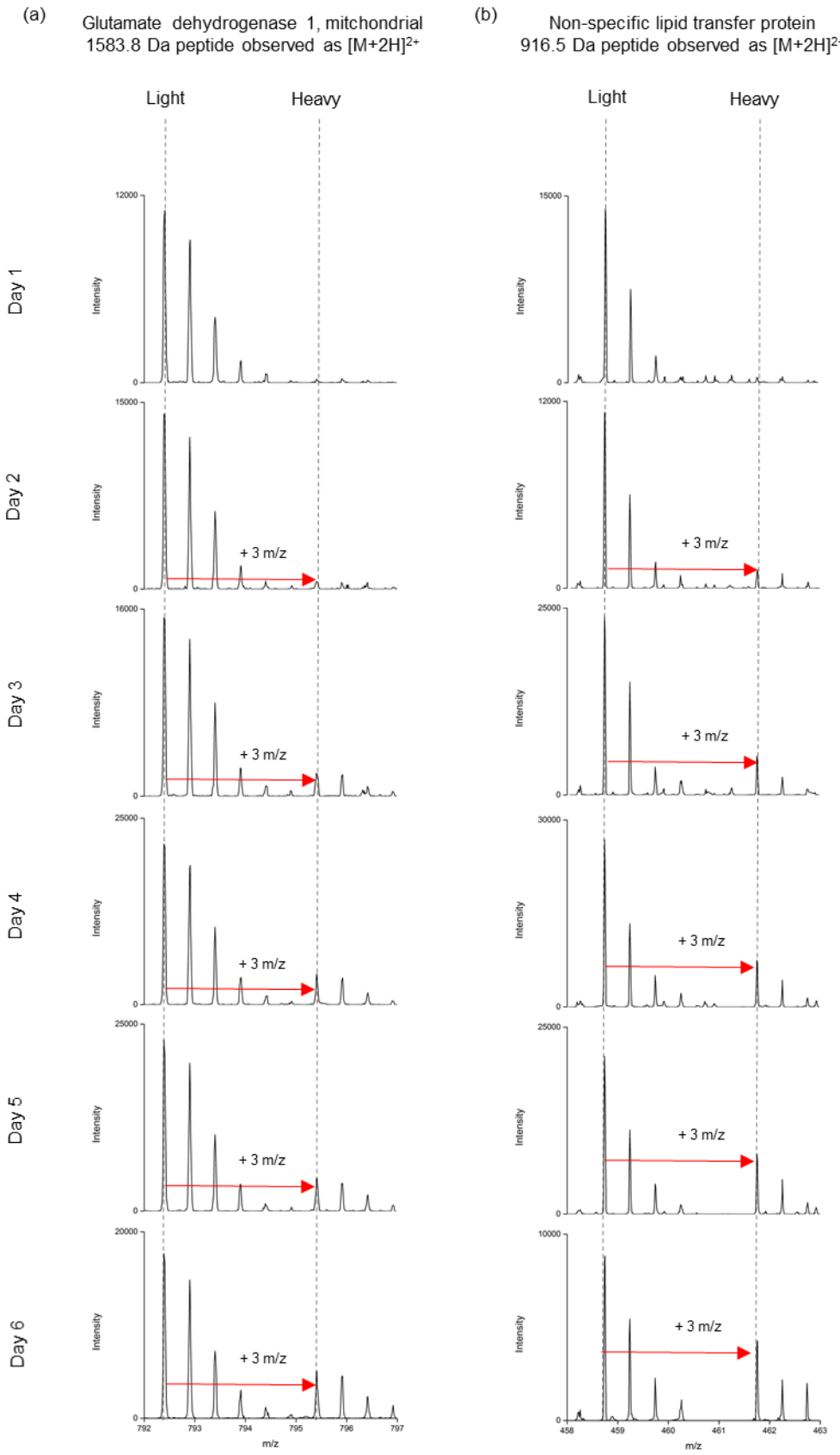
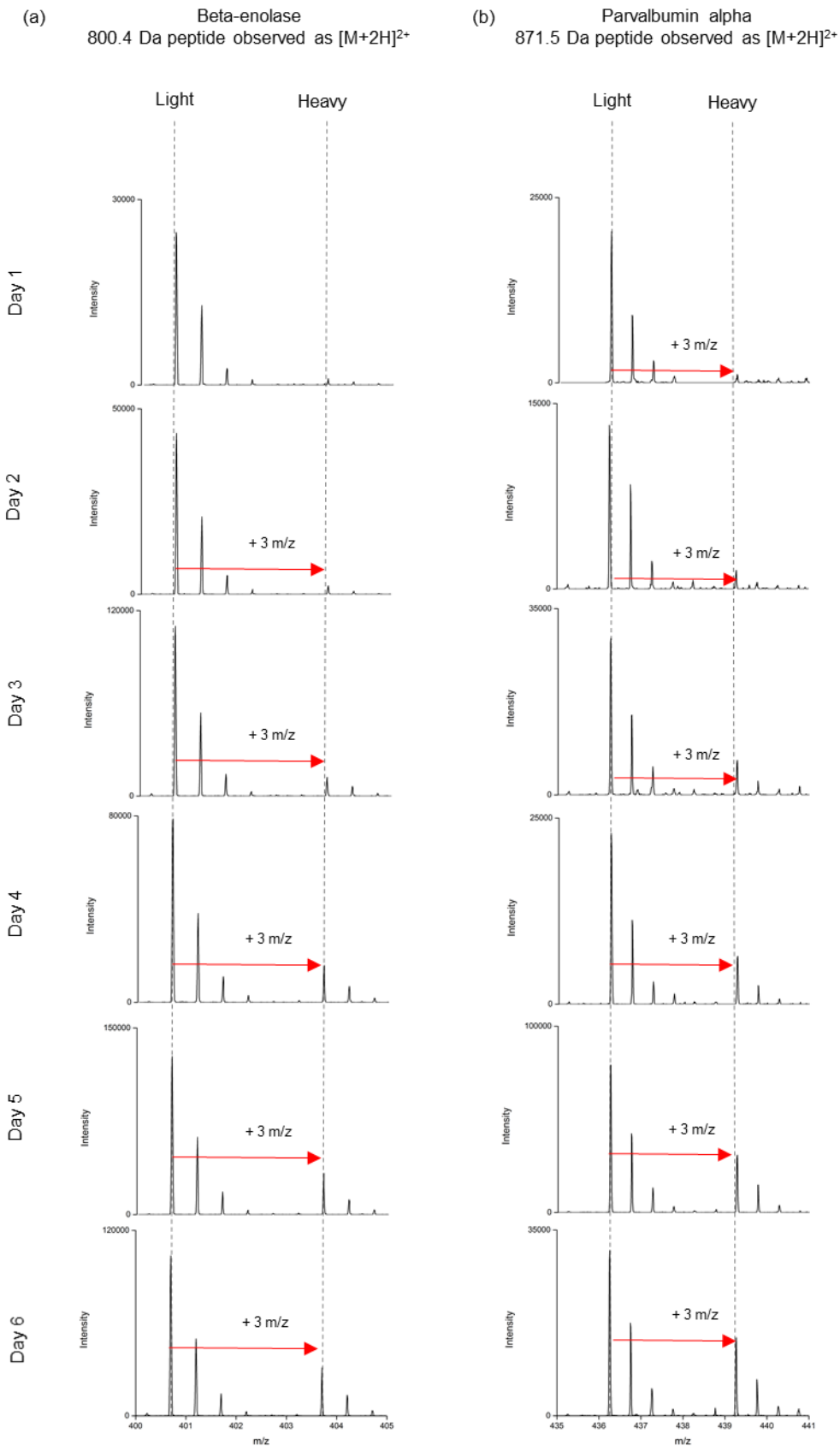

(a)
Glutamate dehydrogenase 1, mitchondrial
1583.8 Da peptide observed as [M+2H]2+
(b)
Non-specific lipid transfer protein
916.5 Da peptide observed as [M+2H]2+
Light
Heavy
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Intensity
+ 3 m/z
m/z

Figure 4.20 Incorporation of [$^{13}C_6$] lysine labelled amino acids over the course of 6 days, shown by the relative abundances of heavy/light peptide profiles in muscle samples from pups. (a) the incorporation of [$^{13}C_6$] lysine into the 800.4 Da peptide of the low turnover protein beta-enolase (b) the incorporation of [$^{13}C_6$] lysine into the 871.5 Da peptide of the high turnover parvalbumin alpha.
Muscle samples were taken from different pups on each day of the six-day experiment. Samples were homogenised, digested in-solution with Lys-C and subjected to LC-MS analysis. Shown are the unlabelled ('light') and labelled ('heavy') peptide profiles of a beta-enolase peptide (a) and a parvalbumin alpha peptide (b), confirming the incorporation of [$^{13}C_6$] lysine into these proteins over the 6-day period.


(a)
Beta-enolase
800.4 Da peptide observed as [M+2H]2+
(b)
Parvalbumin alpha
871.5 Da peptide observed as [M+2H]2+
Light
Heavy
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Intensity
+ 3 m/z
m/z

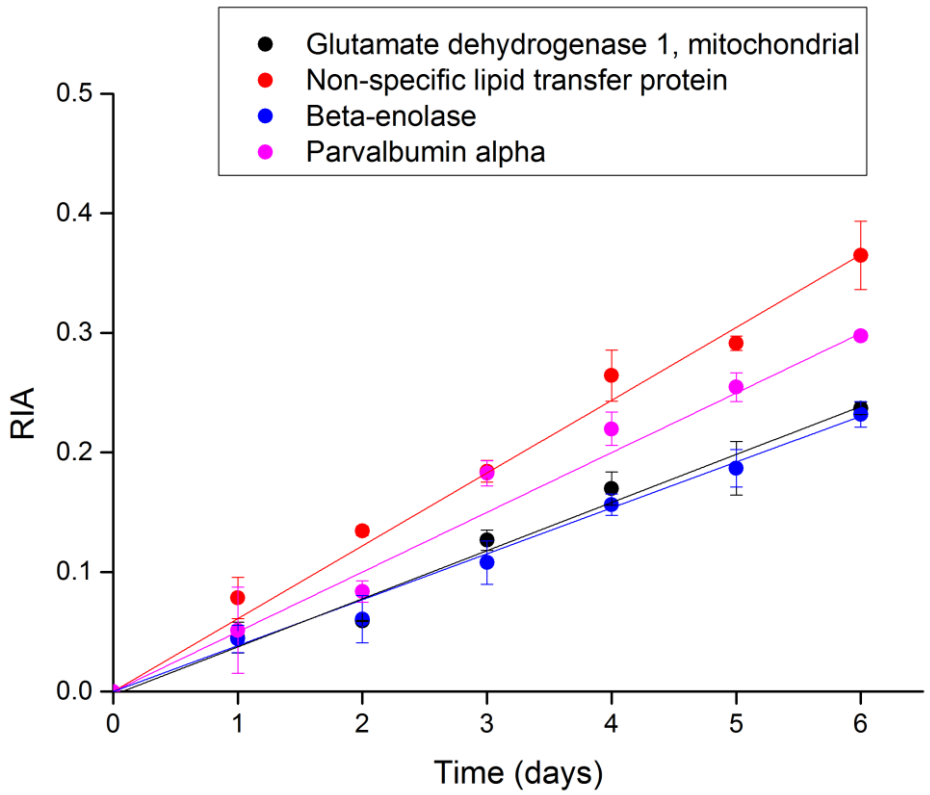


Figure 4.21 The rate of heavy amino acid incorporation over 6 days in the tissue samples of pups whose mothers were fed a [$^{13}C_6$] lysine labelled diet.
Liver and muscle samples were taken from at least one pup each day over the six-day period, homogenised and digested with Lys-C and analysed using LC-MS. The plot shows the average daily RIAs (calculated from the top two peptide matches in each protein, from each animal) for a high and low turnover protein in liver (non-specific lipid transfer protein and glutamate dehydrogenase 1, mitochondrial, respectively), and for a high and low turnover protein in muscle (parvalbumin alpha and beta-enolase, respectively) (error ± SD, n = 4).

will be incorporated into its tissues. The amount of each label present in the pup's tissues will indicate the proportion of investment from each mother. Conversely, if a pup receives investment from only one mother, only that mother's label will be present in the pup's tissues.

4.2.4.1 Experimental set up

The experiment was set up at the Leahurst campus by Dr. J. P. Green (Figure 4.22). Each female was ensured to have a functional RFID tag, and were then paired for familiarisation: eight of the female pairs were related (sisters), and the other nine pairs were unrelated. Each pair were familiarised with each other, the communal cage with feeding mechanism and the unlabelled diet for at least ten days. Each female pair was then mated to two separate males – in related pairs, each female was mated with a male pair where each male was unrelated to the other male and the female pair; in unrelated pairs, each female was mated with a male pair who were siblings but unrelated to the female pair. Mating was staggered within the pairs – for example, the first female of the pair was mated two days before the second female was mated. The reason for this was to aim for the two litters in the communal nest to be born two days apart (no longer than five days apart), allowing easy identification of the two different litters and the number of pups in each based on pup size and skin colour (pup skin colour darkens with age). After mating, each female pair was reintroduced to one another in their communal cages and were left to give birth. When the first litter in a communal nest was 7 days old, the unlabelled diet was removed from the feeding mechanism and replaced with two new labelled diets, one for each female. One of the diets was [D_4] lysine labelled, the other was [D_9] lysine labelled. These labels were used due to their availability; it was deemed more cost effective than to purchase more of the previously used [$^{13}C_6$] lysine and [$^{2}H_8$] valine labels. The females consumed the labelled diet for 7 days whilst nursing their pups. At the end of the 7-day period, urine samples were taken from each female for analysis to ensure each female had fully incorporated the label assigned to it, and had consumed none of the 'wrong' label. At this point, tissue samples (liver and muscle) were taken from each pup for analysis. All samples were stored in Eppendorf tubes and frozen until required.

4.2.4.2 Communal nursing in related female pairs

Protein in the mothers' urine was digested in-solution with Lys-C and analysed using LC-MS. Unlabelled MUP peptides were identified in the mass spectra by searching

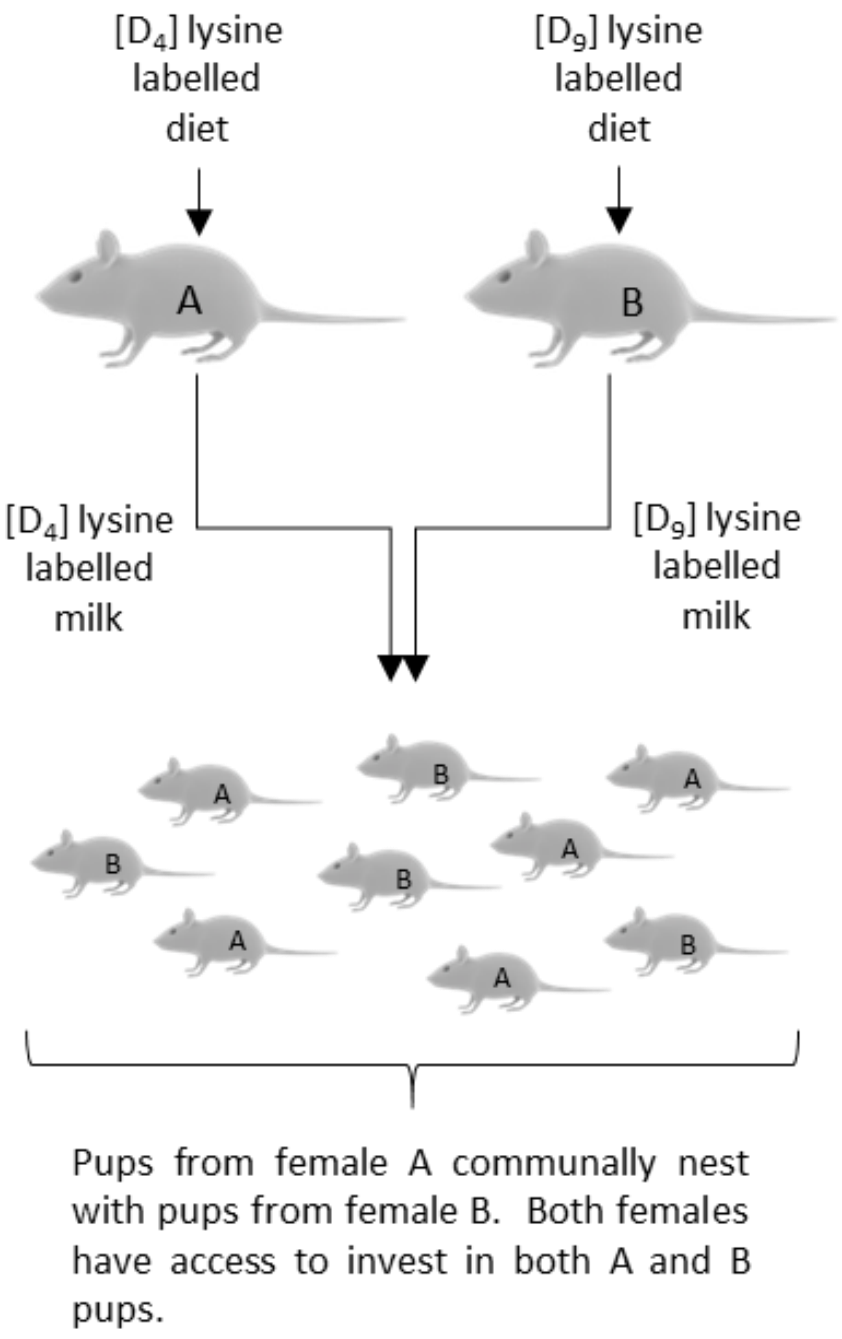


Figure 4.22 The experimental set up at Leahurst to track investment in communal litters.
Pregnant females lived and gave birth in pairs (8 related pairs, 9 unrelated pairs). When both females had given birth, one female was fed [D_4] lysine labelled diet and the other was fed [D_9] lysine labelled diet. These labels were then passed to the pups they invested in via their labelled milk, allowing the investment from each mother to be tracked due to the presence of labels in pup tissue samples.

raw data against a mature MUPs database using PLGS, along with the corresponding [D_4] lysine labels at a +4 Da shift, and the [D_9] lysine labels at a +9 Da shift. Precursor RIAs for three different MUP peptides in each sample were calculated using the formula outlined previously. Since these peptides (in both their unlabelled and labelled forms) were mainly observed as a doubly- charged ions ($[M+2H]^{2+}$) in the raw mass spectrum, the mass shift for the [D_4] lysine labelled peptide was at +2 m/z, and the [D_9] lysine labelled was at +4.5 m/z (Figure 4.23).

Figure 4.24 shows the calculated precursor RIAs for the 1569 Da MUP peptide in each of the urine samples taken from each of the related female pairs on the final day (day 7) of the communal nursing experiment. The precursor RIAs for both the [D_4] lysine labelled and the [D_9] lysine labelled peptide were calculated for each urine sample, to identify whether any female had consumed any of the 'wrong' label (*i.e.* consumed any of the labelled diet that was not assigned to it). The difference between the monoisotopic (M_0) peaks of the two 1569 $[M+2H]^{2+}$ labelled peptides is 2.5 m/z, meaning that the profiles for the two different labelled peptides can be easily distinguished in the mass spectrum (Figure 4.23).

In all related female pairs, the female assigned the [D_4] lysine diet (female A) appeared to have a fully [D_4] lysine labelled precursor pool, with all calculated precursor RIA values being approximately 0.5 (Figure 4.25). None of these females appeared to consume any of the unassigned [D_9] lysine diet, with the low RIA calculations for this protein as a result of noise in the spectrum, as no true isotopologue profile was observed (Figure 4.25). Not all females assigned the [D_9] lysine labelled diet (female B) appeared to have a fully [D_9] lysine labelled precursor pool – only females in pairs 3, 4 and 7 had calculated precursor RIA values of approximately 0.5 (Figure 4.25). Female B in pair 1 had a partially [D_9] lysine labelled precursor pool, with a calculated RIA of 0.26. This female did not appear to consume any of the unassigned [D_4] lysine diet (Figure 4.25). Lower precursor RIA values are likely to be due to the animal eating less of their labelled diet. Female B in pairs 2, 5 and 6 appeared to have a partially [D_9] lysine labelled precursor pool, but the precursor pool for these animals also appeared to be partially [D_4] lysine labelled, with RIA calculations of around 0.1. The mass spectra of the urine from these three mice was studied for the presence of the [D_4] peptide, rather than just spectral noise, using MS isotope to provide a reference of the expected isotope distribution of the 'light' (unlabelled) peptide (Figure 4.24). The female assigned the [D_9] lysine labelled diet in pairs 2 and 5 appeared to have a small amount of the [D_4] lysine labelled peptide

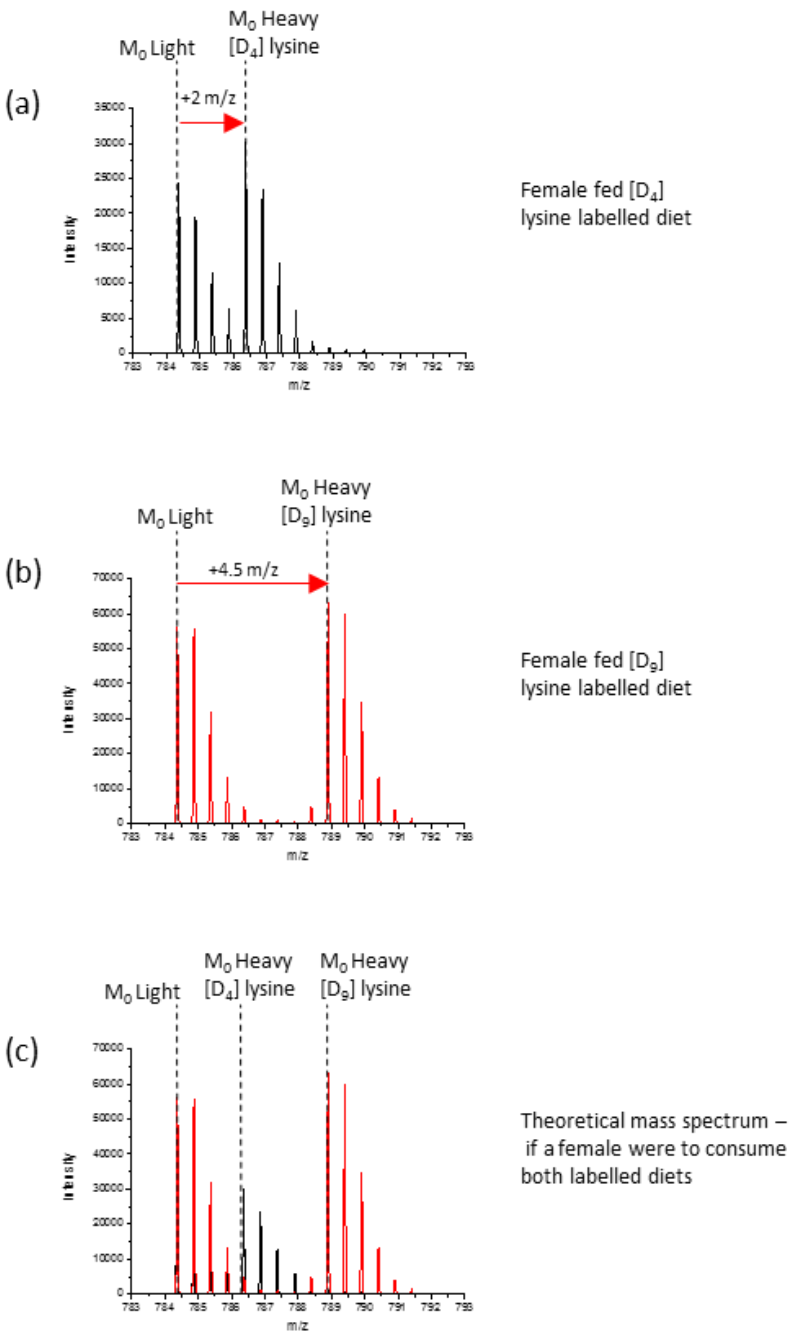
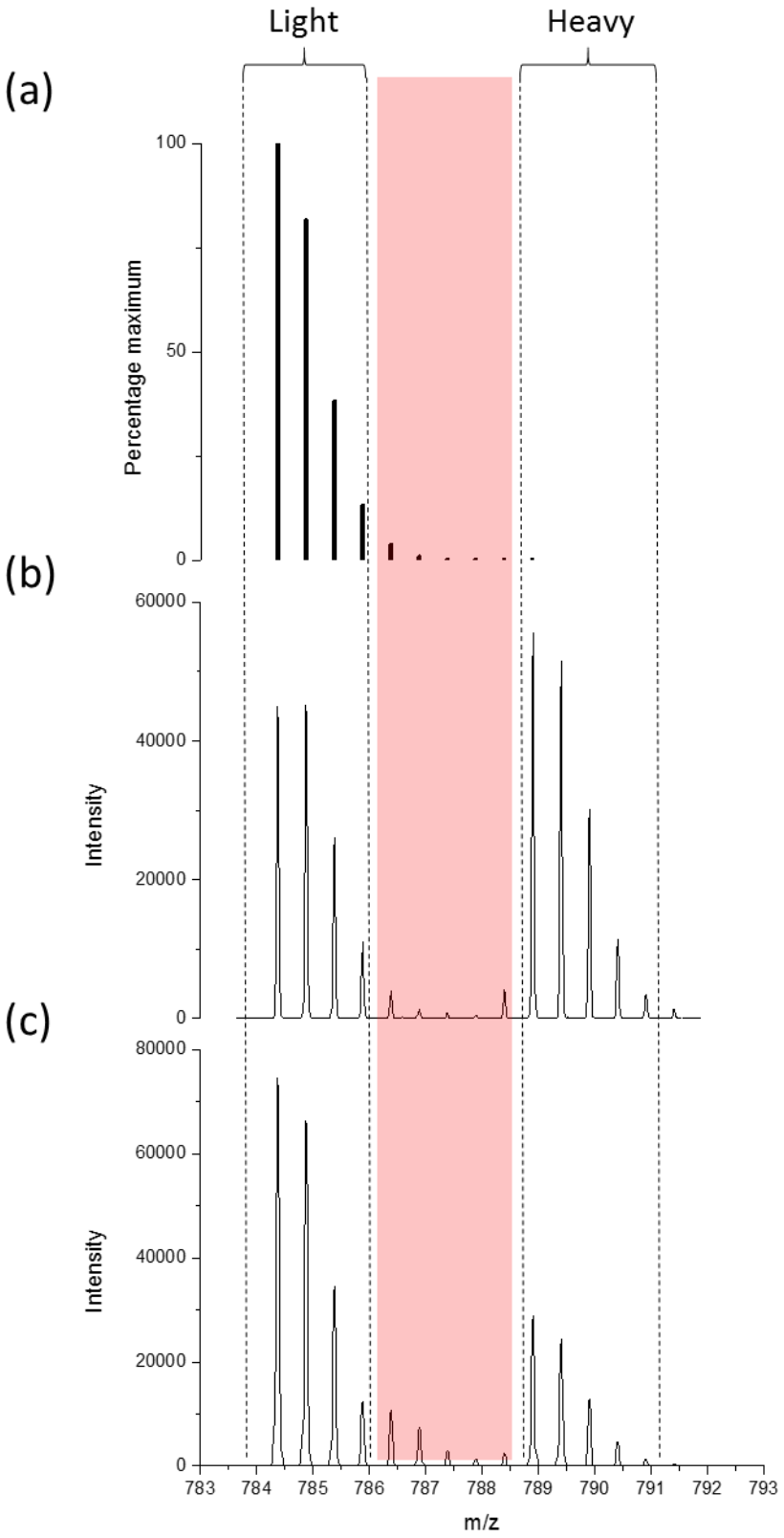


Figure 4.23 The incorporation of (a) [D_4] lysine and (b) [D_9] lysine into the 1567 Da MUP peptide after 7 days of labelling, and (c) the difference between a [D_4] lysine containing (black profile) and [D_9] lysine containing (red profile) 1567 Da MUP peptide.

Urine samples were taken from the females after the 7 day labelling experiment. (a) The labelled ('heavy') [D_4] lysine peptide is 4 Da heavier than the unlabelled ('light') MUP peptide due to the incorporation of D_4 into the lysine residue. The peptide is observed as a $[M+2H]^{2+}$ ion and so the labelled peptide is at +2 m/z. (b) The labelled [D_9] lysine peptide is 9 Da heavier than the unlabelled MUP peptide due to the incorporation of D_9 into the lysine residue. The peptide is observed as a $[M+2H]^{2+}$ ion and so the labelled peptide is at +4.5 m/z. (c) The difference between labelled profiles is key for ascertaining which label is present in the animal. If the feeding mechanism was not working correctly, both 'heavy' labels would be present in the MUP peptides of a single mouse, as shown in the theoretical mass spectrum. This was generated by overlaying the spectra in (a) and (b).

Figure 4.24 (a) Isotope distribution profile of the unlabelled 1569 Da MUP peptide (b) the [D_9] lysine labelled and unlabelled 1569 Da MUP peptide profiles with no [D_4] lysine peptide present, and (c) the [D_9] lysine labelled and unlabelled 1569 Da MUP peptide profiles with [D_4] lysine peptide present.
(a) The isotope distribution profile for the unlabelled ('light') 1569 $[M+2H]^{2+}$ MUP peptide as determined by MS Isotope. This was used to determine whether true [D_4] lysine isotopologues were present in the urine of females assigned the [D_9] lysine diet. (b) Mass spectrum of the unlabelled ('light') and [D_9] lysine ('heavy') 1569 $[M+2H]^{2+}$ MUP peptide, from LC-MS analysis of urine from female B, from pair 1 of related pairs, assigned the [D_9] lysine diet. Shaded in red is the area of the spectrum where a [D_4] lysine profile would be observed if the female had consumed any of the unassigned [D_4] lysine diet. No true isotopologue profile is observed. (c) The same as in (b), however a small but true [D_4] lysine isotopologue profile is present in the spectrum, suggesting this female consumed some of the unassigned [D_4] lysine diet.


Light
Heavy
(a)
100
50
0
Percentage maximum
(b)
60000
40000
20000
0
Intensity
(c)
80000
60000
40000
20000
0
Intensity
783
784
785
786
787
788
789
790
791
792
793
m/z

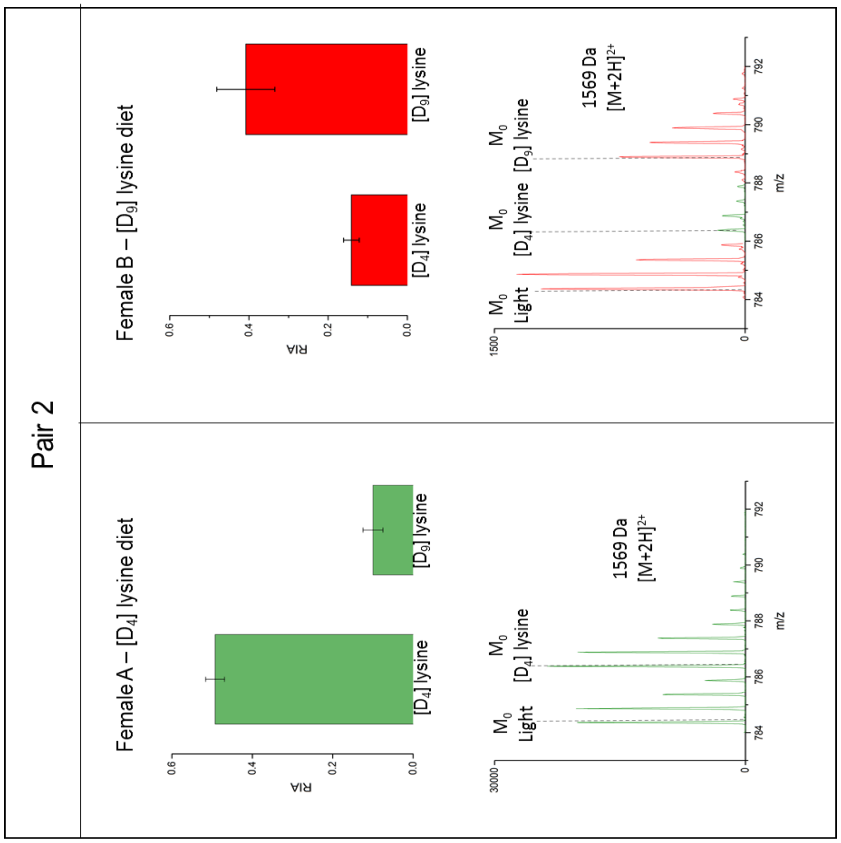
in their urine, as a true isotopologue profile was observed (Figure 4.25). Female B in pair 6 just had a partially labelled [D_9] lysine labelled precursor pool, with no evidence of a true [D_4] lysine peptide in the mass spectrum (Figure 4.25). Since urine samples were only taken on the final day of the experiment, it is unknown whether the females that appeared partially labelled were partially labelled throughout the whole experiment, whether they were less labelled prior to day 7, or possibly fully labelled before reducing their food intake on the final couple of days, resulting in the lower final calculated RIAs. It can be assumed that the fully labelled females reached their maximum precursor RIA after approximately 1 day of being fed the labelled diet, as determined in prior experiments. For this reason, for the assessment of investment in pups, only litters with both fully labelled mothers will be used.

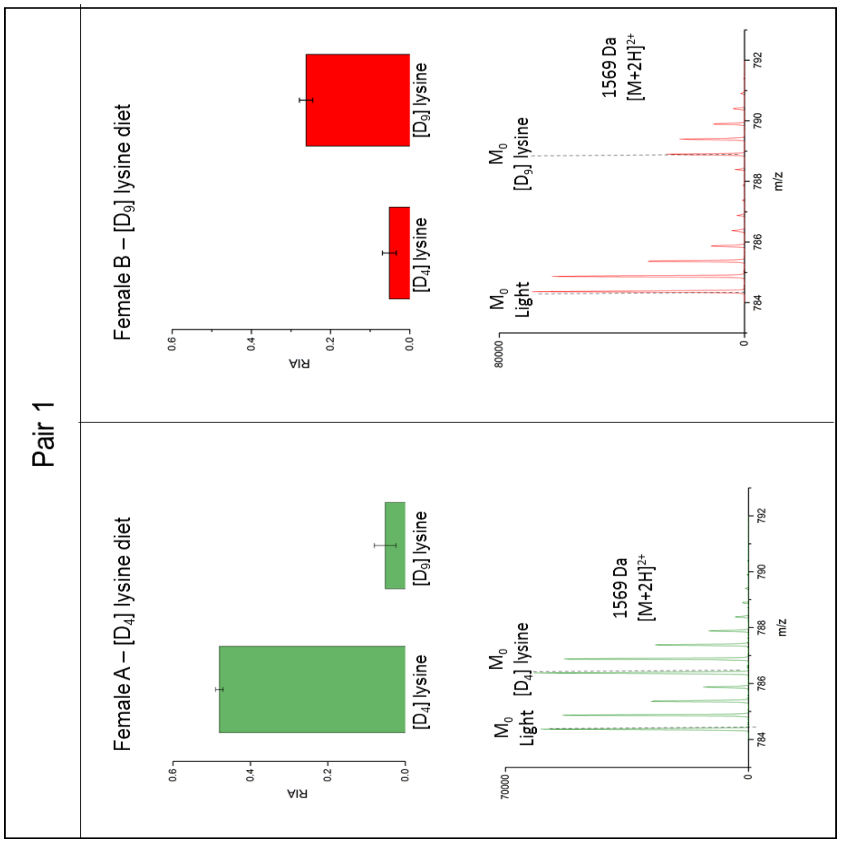
From each pup in each of the three communal litters from pairs 3, 4 and 7, after the 7-day labelling experiment, liver and muscle samples were taken, homogenised, digested with Lys-C and analysed using LC-MS. For each liver sample, precursor RIAs for non-specific lipid transfer protein (high turnover) and glutamate dehydrogenase 1 (low turnover), as determined in the milk labelling pilot study, were calculated using the formula stated previously. For each muscle sample, precursor RIAs for parvalbumin alpha (high turnover) and beta-enolase (low turnover), as determined in the milk labelling pilot study, were also calculated. The differences in turnover rates of these proteins may provide an insight into the investment received by each pup at different points during the 7 day experiment, so will be studied separately. Every pup received investment from both females, with [D_4] lysine and [D_9] lysine labelled peptides present in all pup liver and muscle samples (Figure 4.26).

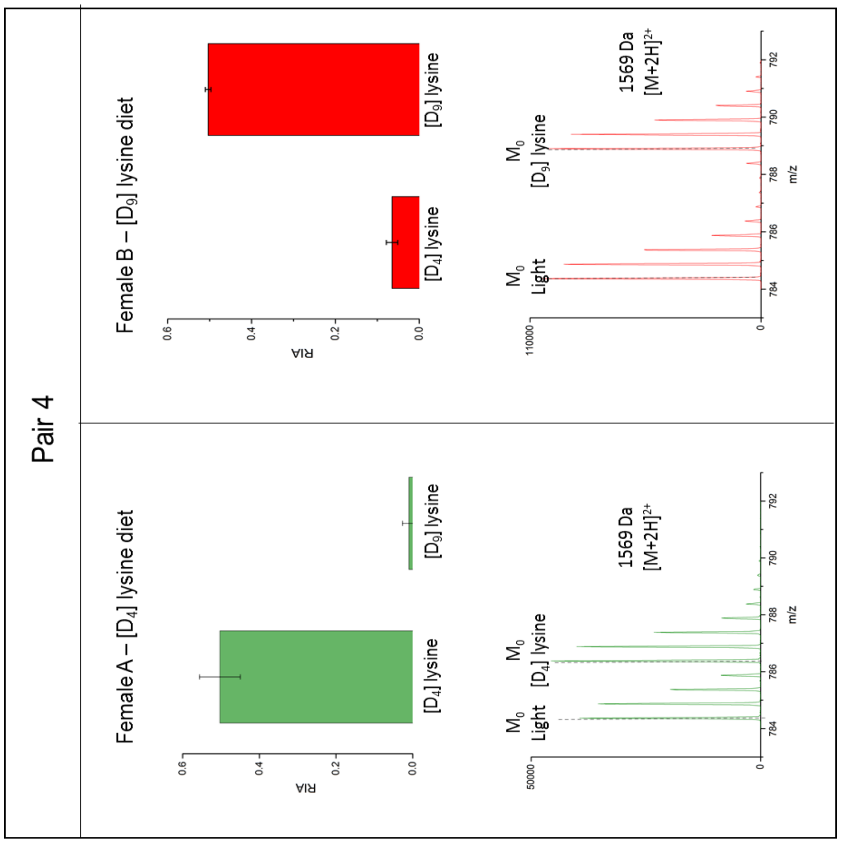
For determining whether there was any statistical significance between the investment from a female in her own pups and in the pups from a related female, the difference between group means and the variation among and between groups was assessed for each of the four proteins using analysis of variance (ANOVA). Where sample size was too small for ANOVA, a Welch two-sample t-test was used; all ANOVA and t-test tables are in the Supplementary Material. Statistical analysis was conducted using R (www.r-project.org), a free software for statistical analysis and graphics. To determine whether either female invested significantly more (or less) in the entire communal litter, the difference between entire group means was assessed using the Welch two sample t-test.

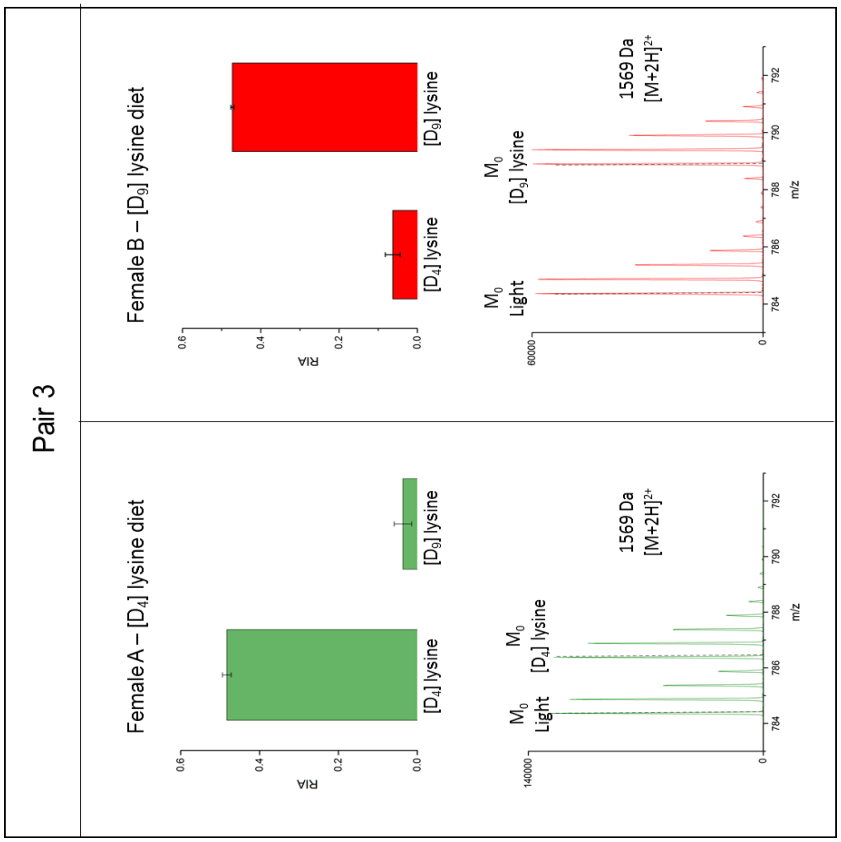
Figure 4.25 The RIA of heavy amino acid incorporation reached after 7 days in pairs of female mice assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet in nests where females were related and both gave birth to litters (pages 197 - 200).

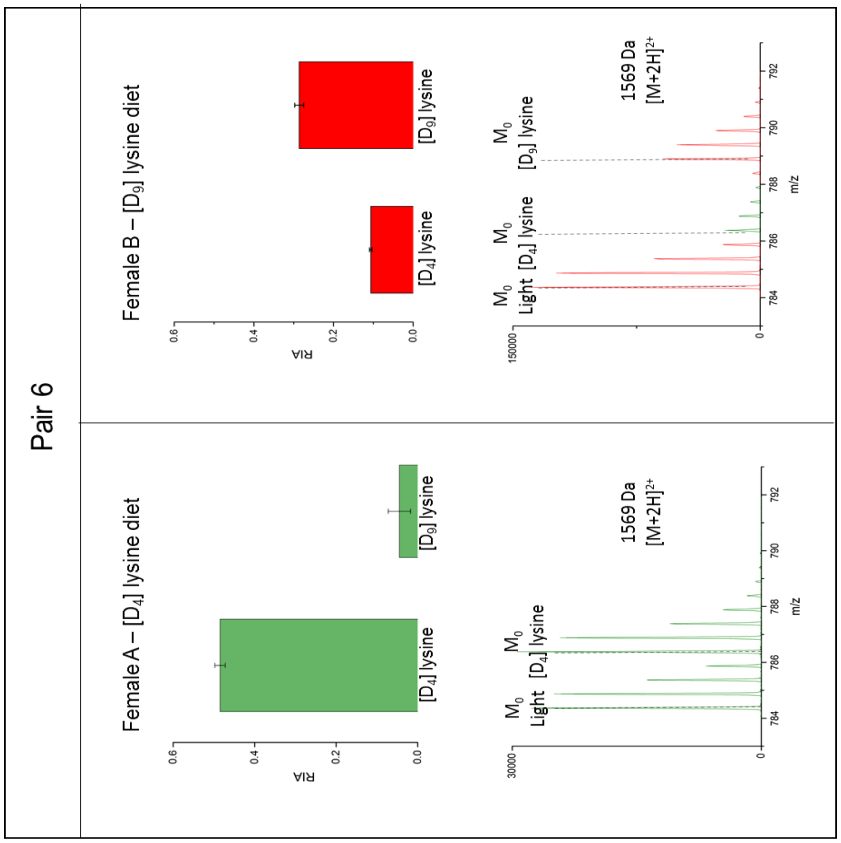
Urine samples were taken from the BALB/c females on the final day of the communal nursing experiment and were subjected to in-solution digestion with Lys-C and analysed using LC-MS. The RIAs for [D_4] lysine and [D_9] lysine were calculated for three MUP peptides in each urine sample (error ± SD, n = 3). The isotope distribution patterns for the 1569 Da $[M+2H]^{2+}$ MUP peptide are shown for each female urine sample – the [D_4] lysine peptides are shown in green and the [D_9] lysine peptides are shown in red. The unlabelled peptides are shown in green for females assigned the [D_4] lysine diet, and red for females assigned the [D_9] lysine labelled diet.

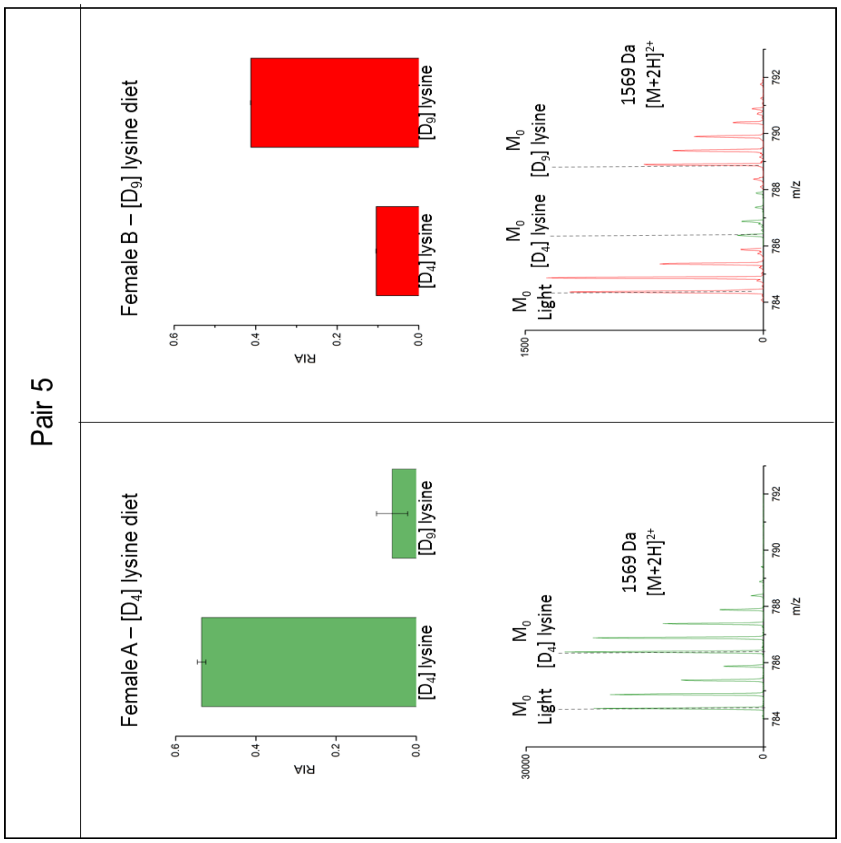

Pair 2
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0
Light
[D4] lysine
[D9] lysine
1569 Da
[M+2H]2+
m/z

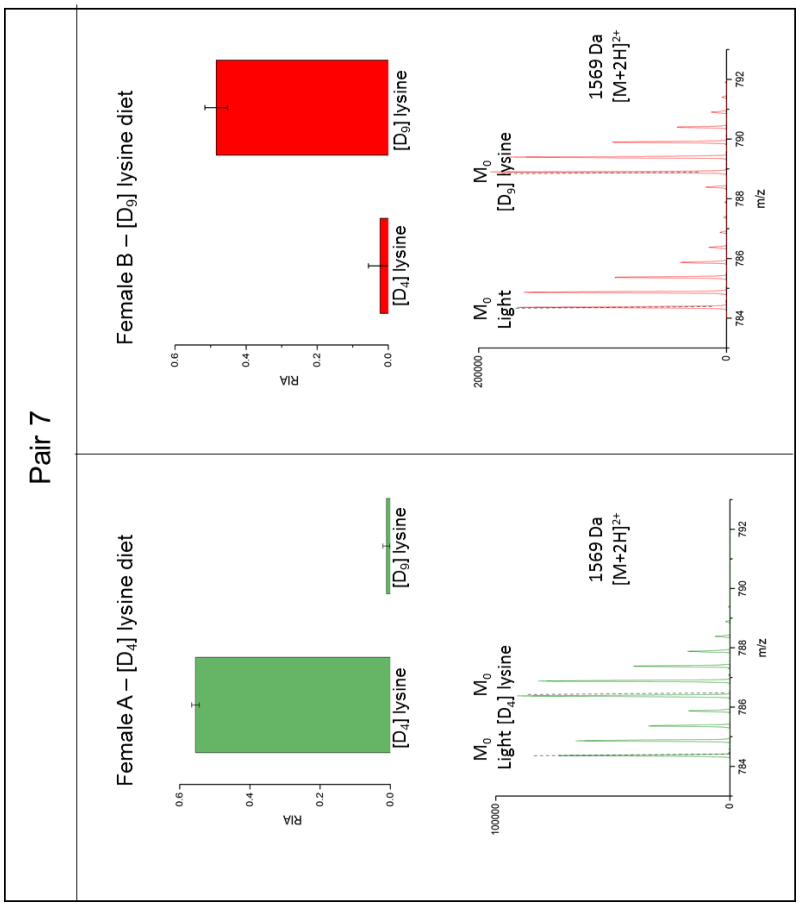

Pair 1
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0
Light
[D4] lysine
[D9] lysine
1569 Da
[M+2H]2+
m/z


Pair 4
Female B – [D9] lysine diet
[D9] lysine
[D4] lysine
RIA
M0 Light
M0 [D9] lysine
1569 Da [M+2H]2+
m/z
Female A – [D4] lysine diet
[D9] lysine
[D4] lysine
RIA
M0 Light
M0 [D4] lysine
1569 Da [M+2H]2+
m/z


Pair 3
Female B – [D9] lysine diet
[D9] lysine
[D4] lysine
RIA
M0 Light
M0 [D9] lysine
1569 Da [M+2H]2+
m/z
Female A – [D4] lysine diet
[D9] lysine
[D4] lysine
RIA
M0 Light
M0 [D4] lysine
1569 Da [M+2H]2+
m/z


Pair 6
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D9] lysine
[D4] lysine
M0
Light
[D4] lysine
[D9] lysine
1569 Da
[M+2H]2+
m/z


Pair 5
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D9] lysine
[D4] lysine
M0
Light
[D4] lysine
[D9] lysine
1569 Da
[M+2H]2+
m/z


Pair 7
Female B – [D9] lysine diet
[D9] lysine
[D4] lysine
RIA
0.6
0.4
0.2
0.0
M0
[D9] lysine
M0
Light
1569 Da
[M+2H]2+
200000
0
784
786
788
790
792
m/z
Female A – [D4] lysine diet
[D9] lysine
[D4] lysine
RIA
0.6
0.4
0.2
0.0
M0
Light [D4] lysine
M0
1569 Da
[M+2H]2+
100000
0
784
786
788
790
792
m/z

Litter 3 (from pair 3)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.27). A and B litter sizes were similar (4 pups and 3 pups, respectively), with litter B being 3 days older than litter A.

The [D_9] lysine RIAs for each protein are higher than the [D_4] lysine RIAs, suggesting that female B consistently invests slightly more than female A in the communal litter. This difference was significant (Welch two sample t-test: t = -2.7058, df = 52.363, p = 0.009176). There is no significant difference between the [D_4] lysine precursor RIAs in A and B pups (ANOVA – glutamate dehydrogenase 1: P =0 .93; beta-enolase: P = 0.585; parvalbumin alpha: P = 0.821; non-specific lipid transfer protein: P = 0.904). There is no significant difference between the [D_9] lysine precursor RIAs in A and B pups (ANOVA – glutamate dehydrogenase 1: P = 0.202; beta-enolase: P = 0.479; parvalbumin alph: P = 0.219; non-specific lipid transfer protein: P = 0.131). The statistical analysis of calculated RIA values indicate neither female invested significantly differently in either litter in the communal nest, but female B appeared to invest significantly more in all pups than female A.

Litter 4 (from pair 4)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.28). A and B litter sizes were 6 pups and 9 pups, respectively, with litter B being 2 days older than litter A.

Again, the [D_9] lysine RIAs for each protein are significantly higher than the [D_4] lysine RIAs, indicating that female B consistently invests slightly more than female A in the communal litter (Welch two sample t-test: t = -3.5129, df = 117.3, p = 0.0006304). There is no significant difference between the [D_4] lysine precursor RIAs in A and B pups (ANOVA – glutamate dehydrogenase 1: P = 0.66; beta-enolase: P = 0.149; parvalbumin alpha: P = 0.57; non-specific lipid transfer protein: P = 0.895) or the [D_9] lysine precursor RIAs in A and B pups (ANOVA – glutamate dehydrogenase 1: P = 0.901; beta-enolase: P = 0.933; parvalbumin alpha: P = 0.932; non-specific lipid transfer protein: P = 0.93). Again, neither female appeared to invest significantly differently in either litter in the communal nest, but female B appeared to invest more in all pups than female A.

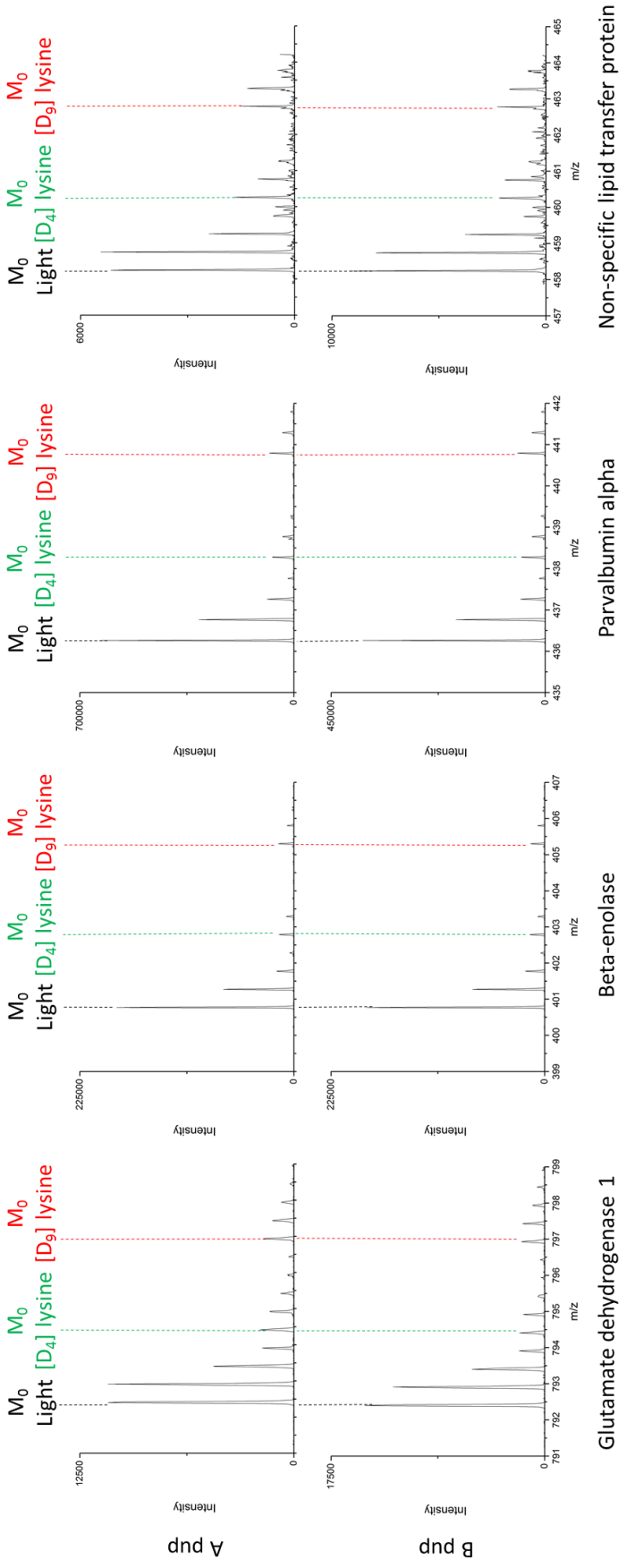


Figure 4.26 Examples of the heavy amino acid incorporation reached after 7 days in the high and low turnover liver and muscle proteins in pups as a result of investment from related mothers who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet. Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. Shown are the unlabelled ('light') and labelled ('heavy') peptide profiles of a glutamate dehydrogenase 1 peptide (low turnover in liver), a beta-enolase peptide (low turnover in muscle), a parvalbumin alpha peptide (high turnover in muscle), and a non-specific lipid transfer protein peptide (high turnover in liver), confirming the incorporation of [D_4] lysine (profiles labelled green) and [D_9] lysine (red) into pup tissue samples as a result of investment from both females. This figure shows the spectra of tissue samples taken from an 'A pup' (whose mother is female A) and a 'B pup' (whose mother is female B) from Litter 3.

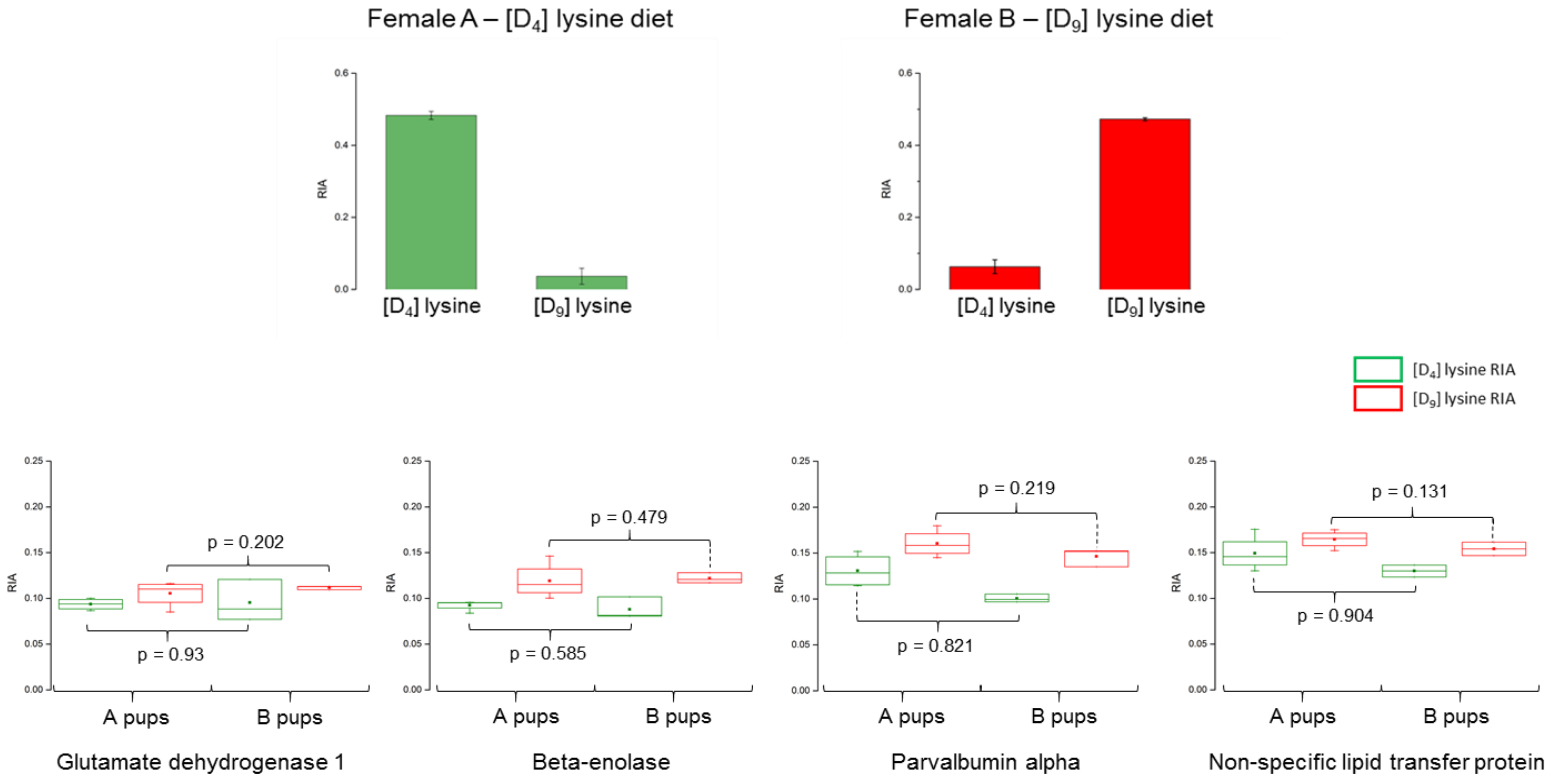


Figure 4.27 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 3 pup liver and pup muscle samples as a result of investment from related mothers (Pair 3), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

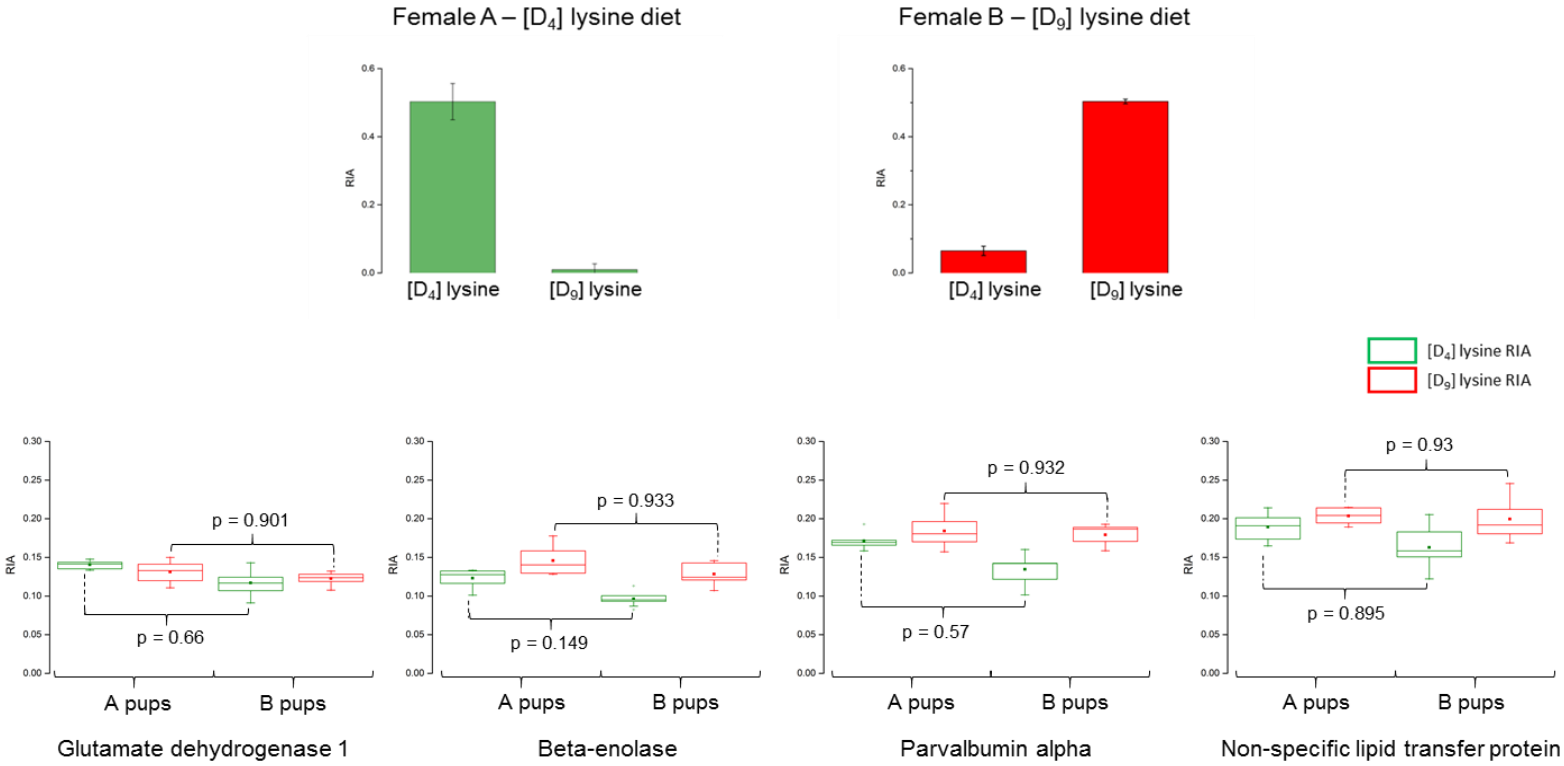


Figure 4.28 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 4 pup liver and pup muscle samples as a result of investment from related mothers (Pair 4), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

Litter 7 (from pair 7)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.29). A and B litter sizes were similar (4 pups and 5 pups, respectively), with litter B being 1 day older than litter A. The [D_4] lysine RIAs for each protein are significantly higher than the [D_9] lysine RIAs, indicating that female A consistently invests more than female B in the communal litter (Welch two sample t-test: t = 7.6728, df = 67.717, p = 8.81e-11). Again, there is no significant difference between the [D_4] lysine precursor RIAs in A and B pups (ANOVA –glutamate dehydrogenase 1: P = 0.157; beta-enolase: P = 0.931; parvalbumin alpha: P = 0.119; non-specific lipid transfer protein: P = 0.718) or the [D_9] lysine precursor RIAs in A and B pups (ANOVA –glutamate dehydrogenase 1: P = 0.194; beta-enolase: P = 0.211; parvalbumin alpha: P = 0.108; non-specific lipid transfer protein: P = 0.214). Again, neither female appeared to invest significantly differently in either litter in the communal nest, but in this nest, female A appeared to invest more in all pups than female B.

In all three communal litters, no female appeared to discriminate between their own pups and their sister's pups in terms of investment. However, in all litters, one female appeared to invest more in the entire communal litter than the other. Due to the small sample size, no reason for this could be determined (*e.g*. did females that gave birth first invest more? Did females with smaller litters invest more? *etc*.).

4.2.4.3 Communal nursing in unrelated female pairs

Protein in the mothers' urine was digested in-solution with Lys-C and analysed using LC-MS, and all analysis and calculations were the same as those for the related female pairs, described in the previous section. In all six unrelated female pairs, the female assigned the [D_4] lysine diet (female A) appeared to have a fully [D_4] lysine labelled precursor pool, with all calculated precursor RIA values being approximately 0.5 (Figure 4.30). None of these females appeared to consume any of the unassigned [D_9] lysine diet, with the low RIA calculations for this protein as a result of noise in the spectrum, as no true isotopologue profile was observed. In all pairs, the female assigned the [D_9] lysine diet (female B) appeared to have a fully [D_9] lysine labelled precursor pool. None of these females appeared to consume any of the unassigned [D_4] lysine diet, with the low RIA calculations for this protein as a result of the unlabelled peptide ion distribution (Figure 4.30).

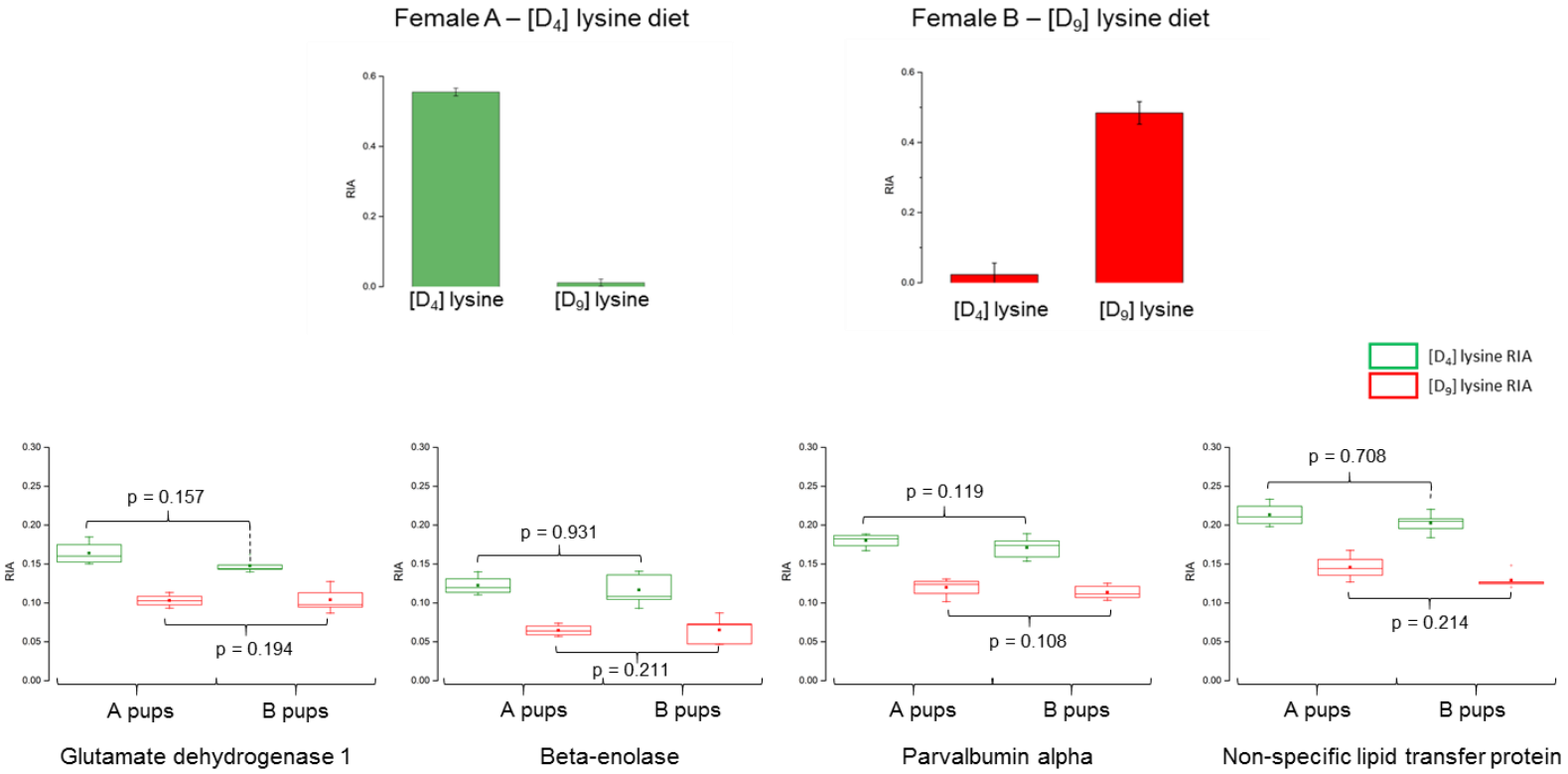
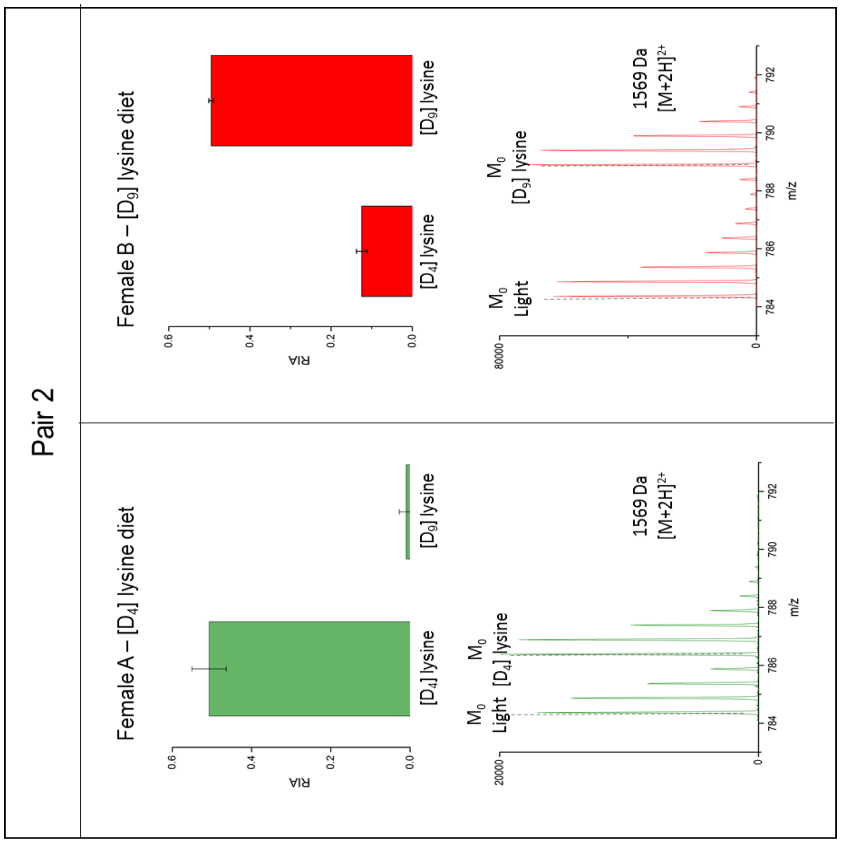


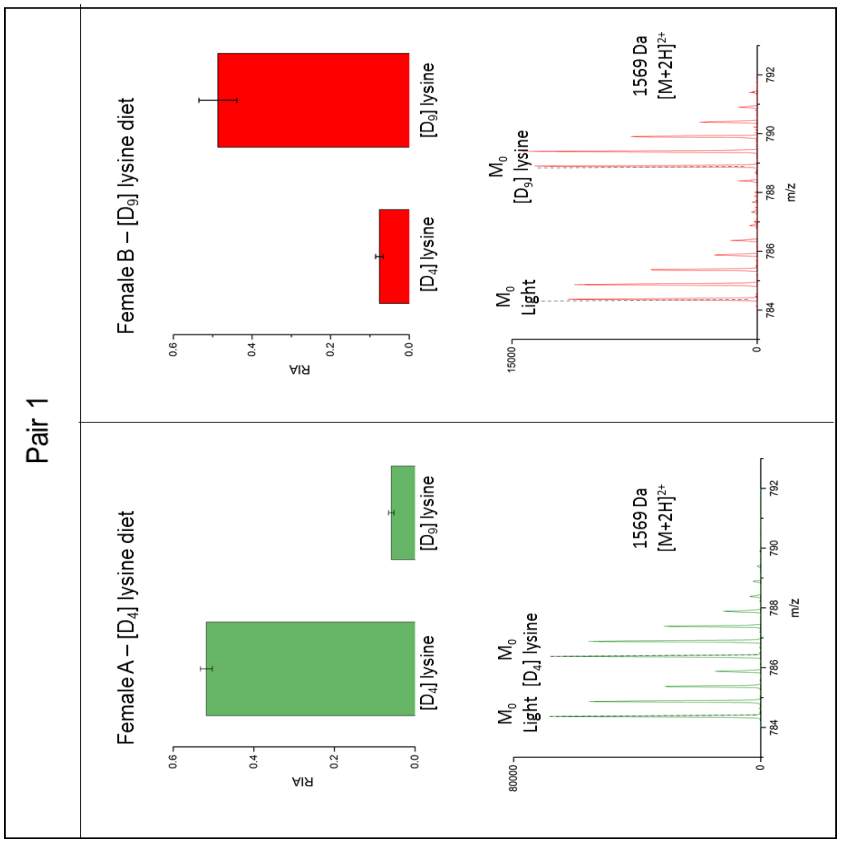
Figure 4.29 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 7 pup liver and pup muscle samples as a result of investment from related mothers (Pair 7), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

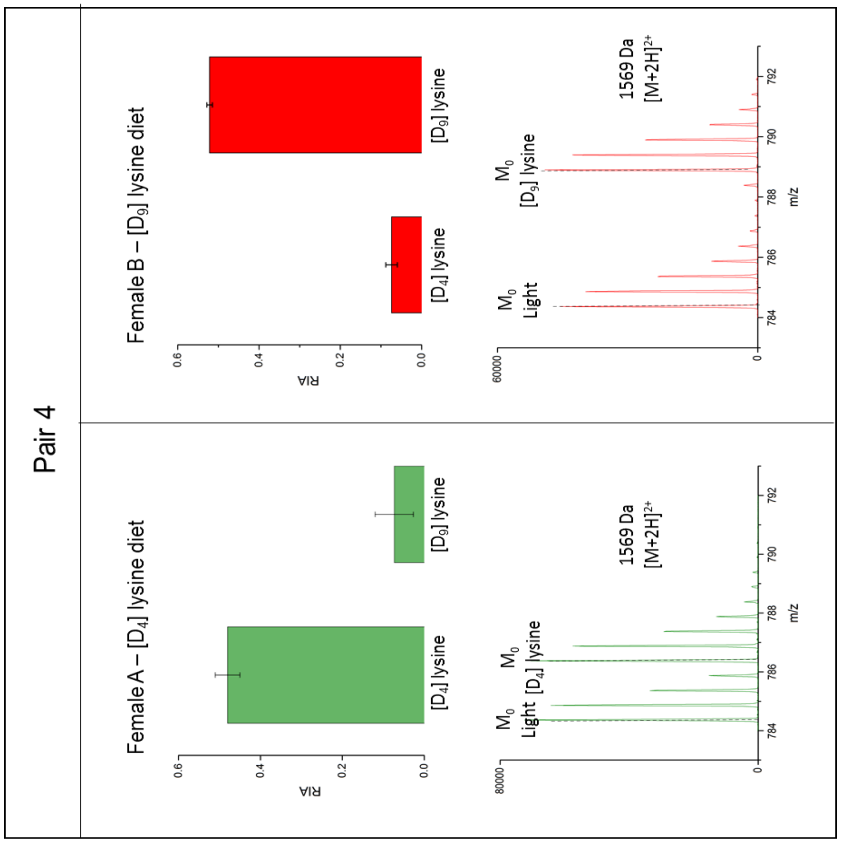
Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

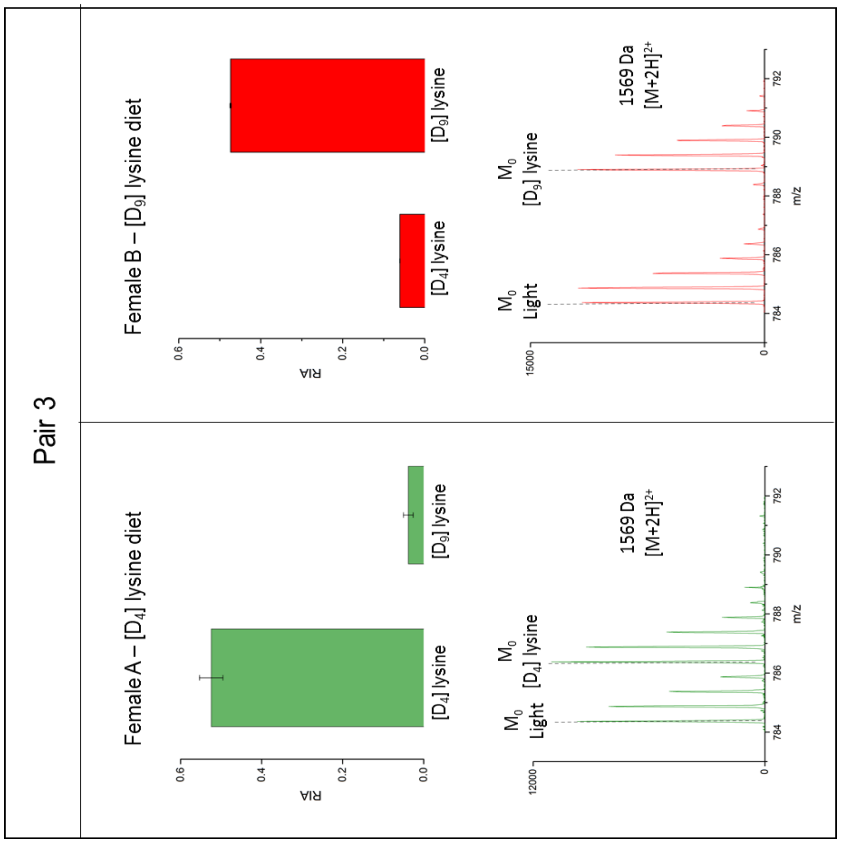
Figure 4.30 The RIA of heavy amino acid incorporation reached after 7 days in pairs of female mice assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet in nests where females were unrelated and both gave birth to litters (pages 208 – 210).

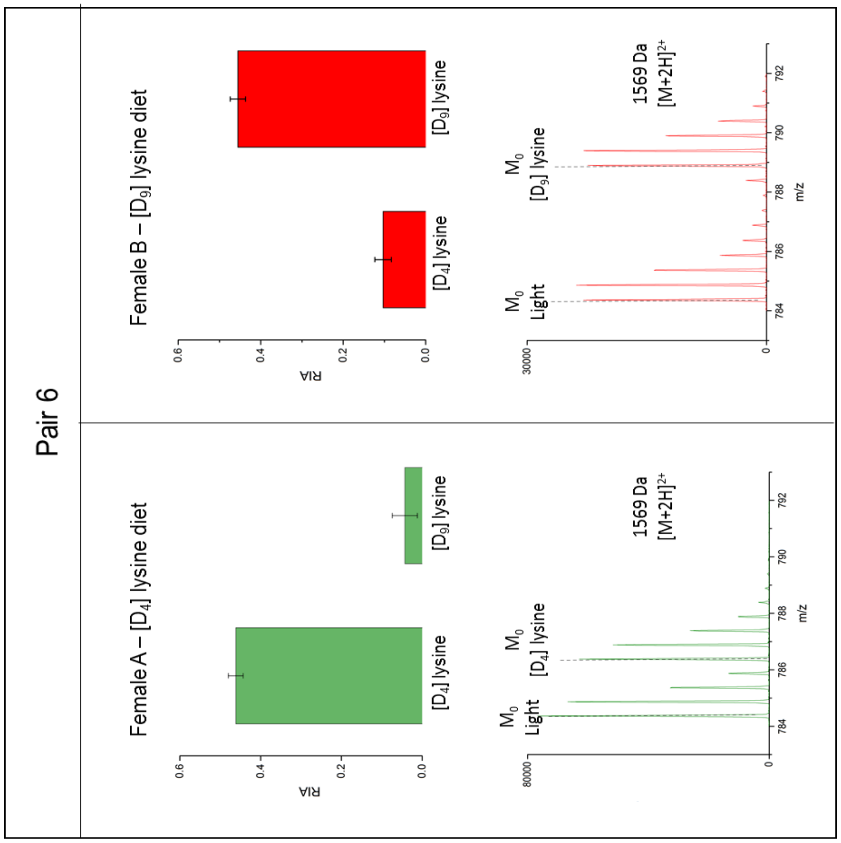
Urine samples were taken from the BALB/c females on the final day of the communal nursing experiment and were subjected to in-solution digestion with Lys-C and analysed using LC-MS. The RIAs for [D_4] lysine and [D_9] lysine were calculated for three MUP peptides in each urine sample (error ± SD, n = 3). The isotope distribution patterns for the 1569 Da $[M+2H]^{2+}$ MUP peptide are shown for each female urine sample – the [D_4] lysine peptides are shown in green and the [D_9] lysine peptides are shown in red. The unlabelled peptides are shown in green for females assigned the [D_4] lysine diet, and red for females assigned the [D_9] lysine labelled diet.

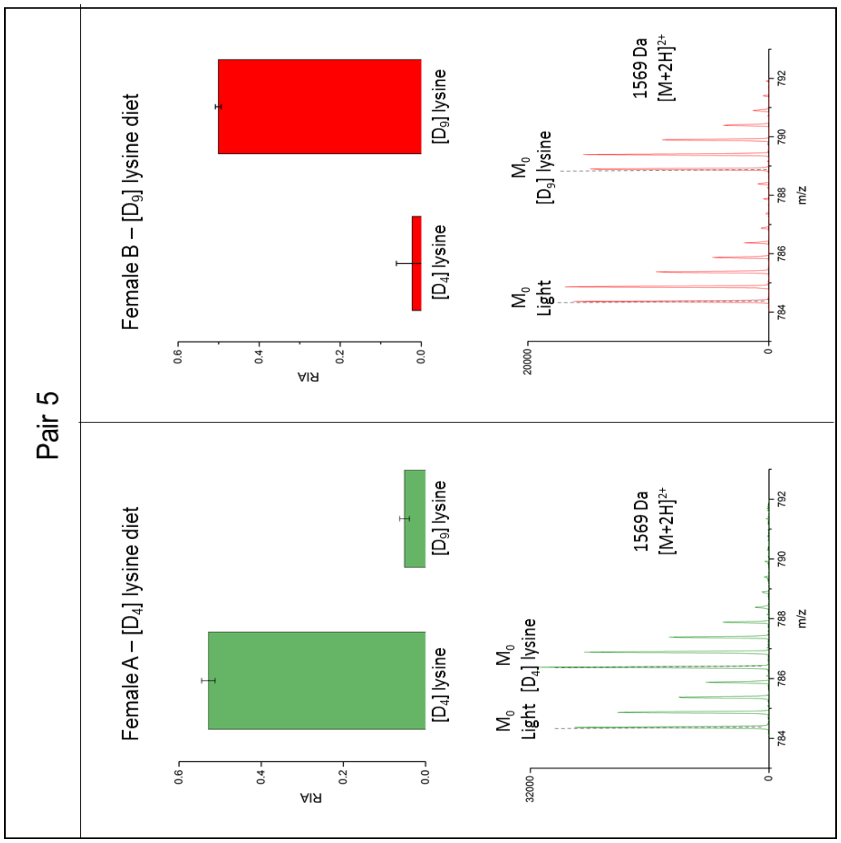

Pair 2
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da [M+2H]2+
m/z


Pair 1
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da [M+2H]2+
m/z


Pair 4
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da
[M+2H]2+
m/z


Pair 3
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da
[M+2H]2+
m/z


Pair 6
Female A – [D4] lysine diet
Female B – [D9] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da
[M+2H]2+
m/z


Pair 5
Female A – [D4] lysine diet
Female B – [D9] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da
[M+2H]2+
m/z

From each pup in each of the six communal litters from pairs 1 - 6, after the 7-day labelling experiment, liver and muscle samples were taken, homogenised, digested with Lys-C and analysed using LC-MS. All analysis was carried out as described in the previous section, and again, every pup received investment from both females, with [D_4] lysine and [D_9] lysine labelled peptides present in all pup liver and muscle samples (Figure 4.31).

Litter 1 (from pair 1)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.32). A and B litter sizes were notably different (7 pups and 2 pups, respectively), with litter B being 1 days older than litter A.

The [D_9] lysine RIAs for each protein are higher than the [D_4] lysine RIAs, suggesting that female B consistently invests slightly more than female A in the communal litter; however, this difference is not significant (Welch two sample t-test: t = -1.74, df = 74.72, p = 0.0856). There is no significant difference between the [D_4] lysine precursor RIAs in A and B pups (Welch two sample t-test – glutamate dehydrogenase 1: P = 0.93; beta-enolase: P = 0.585; parvalbumin alpha: P = 0.821; non-specific lipid transfer protein: P = 0.904). There is no significant difference between [D_9] lysine precursor RIAs in A and B pups (Welch two sample t-test – glutamate dehydrogenase 1: P = 0.202; beta-enolase: P = 0.479; parvalbumin alpha: P = 0.219; non-specific lipid transfer protein: P = 0.131). Neither female invested significantly differently in either litter in the communal nest or in the nest as a whole.

Litter 2 (from pair 2)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.33). A and B litter sizes were notably different (2 pups and 7 pups, respectively), with both litters born on the same day.

There was no significant difference between the [D_4] and [D_9] lysine RIAs in each protein, suggesting both females invested equally in the communal litter (Welch two sample t-test: t = -1.3797, df = 69.352, p = 0.1721). Again, there is no significant difference between the [D_4] lysine precursor RIAs in A and B pups (Welch two sample t-test – glutamate dehydrogenase 1: P = 0.364; beta-enolase: P = 0.058; parvalbumin

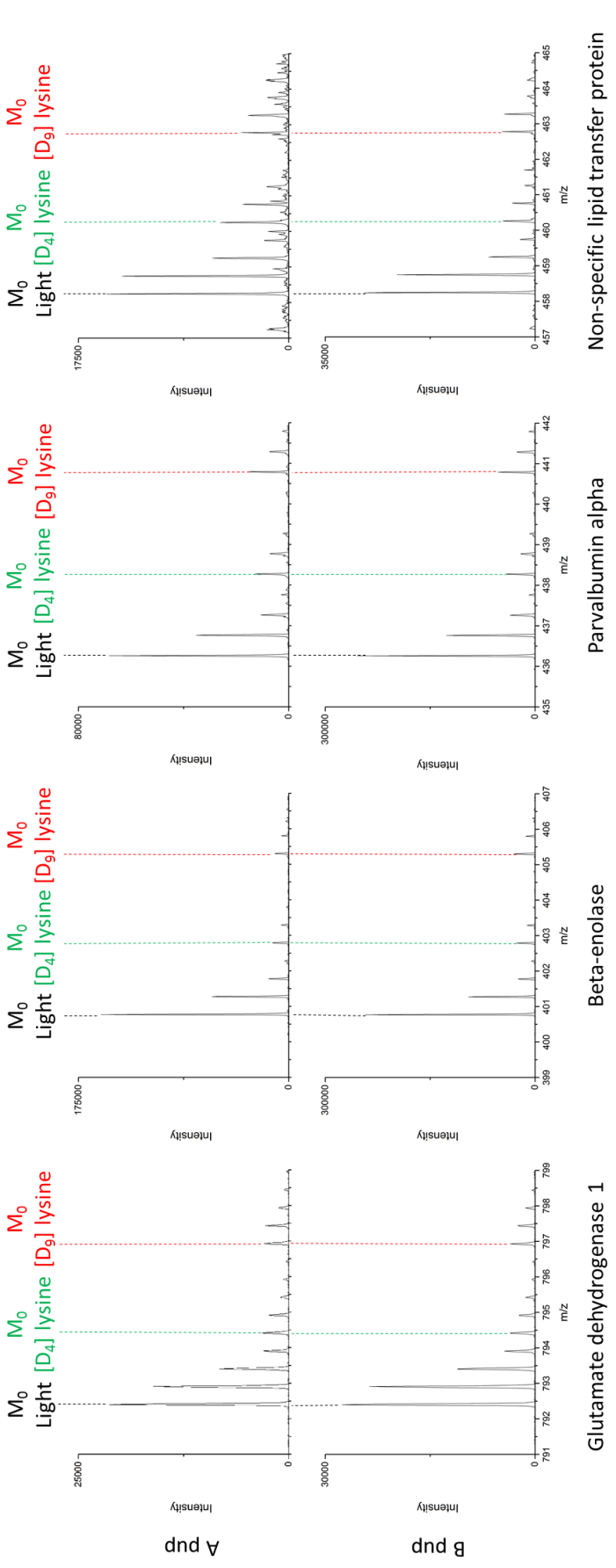


Figure 4.31 Examples of the heavy amino acid incorporation reached after 7 days in the high and low turnover liver and muscle proteins in pups as a result of investment from unrelated mothers who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet. Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. Shown are the unlabelled ('light') and labelled ('heavy') peptide profiles of a glutamate dehydrogenase 1 peptide (low turnover in liver), a beta-enolase peptide (low turnover in muscle), a parvalbumin alpha peptide (high turnover in muscle), and a non-specific lipid transfer protein peptide (high turnover in liver), confirming the incorporation of [D_4] lysine (profiles labelled green) and [D_9] lysine (red) into pup tissue samples as a result of investment from both females. This figure shows the spectra of tissue samples taken from an 'A pup' (whose mother is female A) and a 'B pup' (whose mother is female B) from Litter 1.

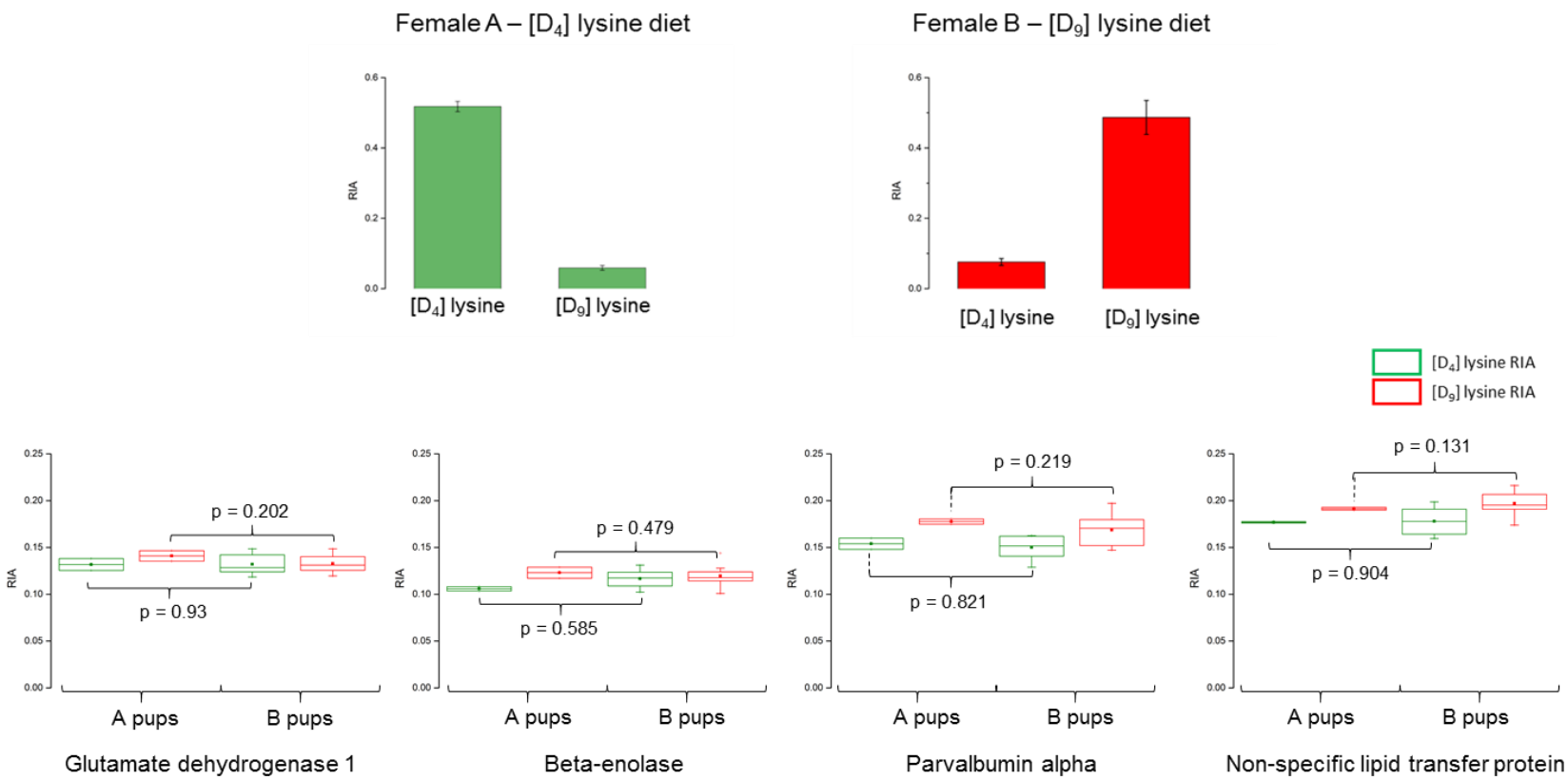


Figure 4.32 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 1 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 1), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

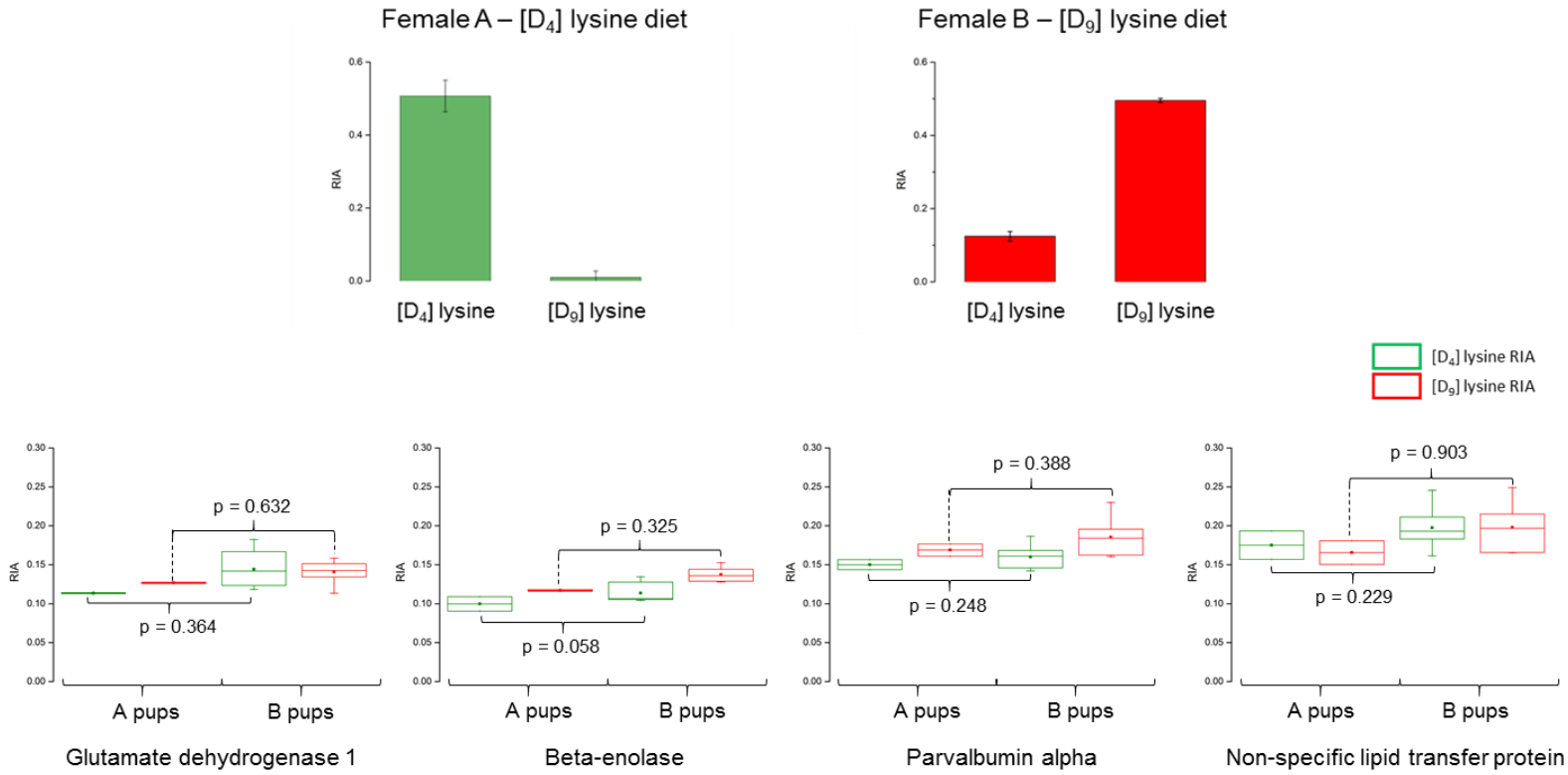


Figure 4.33 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 2 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 2), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

alpha: P = 0.248; non-specific lipid transfer protein: P = 0.229). There is also no significant difference between the [D_9] lysine precursor RIAs in A and B pups (Welch two sample t-test – glutamate dehydrogenase 1: P = 0.632; beta-enolase: P = 0.325; parvalbumin alpha: P = 0.388; non-specific lipid transfer protein: P = 0.903). Neither female appeared to invest significantly differently in either litter in the communal nest, and neither female appeared to invest more (or less) than the other in the communal litter.

Litter 3 (from pair 3)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.34). A and B litter sizes were similar (7 pups and 8 pups, respectively), with litter A born 4 days prior to litter B.

There was no significant difference between the [D_4] and [D_9] lysine RIAs in each protein, suggesting both females invested equally in the communal litter (Welch two sample t-test: t = 0.4614, df = 113.197, p = 0.6454). There is no significant difference between the [D_4] (ANOVA: P = 0.38) and [D_9] lysine (ANOVA: P = 0.344) precursor RIAs in A and B pups in the low turnover liver protein (glutamate dehydrogenase 1), however in the low turnover muscle protein (beta-enolase), there appeared to be significantly higher [D_9] lysine RIA (from female B) in her own litter (litter B) than litter A (ANOVA: P = 0.066), but there is no significant difference between the [D_4] lysine precursor RIAs in A and B pups (ANOVA: P = 0.614). In the high turnover muscle protein (parvalbumin alpha), there appeared to be significantly higher [D_9] lysine RIA (from female B) in her own litter (litter B) than litter A (ANOVA: P = 0.079), but again, no significant difference between the [D_4] lysine precursor RIAs in A and B pups (ANOVA: P = 0.82). In the high turnover liver protein (non-specific lipid transfer protein), there is no significant difference between the [D_4] (ANOVA: P = 0.945) and [D_9] lysine (ANOVA: P = 0.164) precursor RIAs in A and B pups. Although these differences have been deemed significant, they are not highly significant (all P values are <0.05, but >0.01). This means there is between 95 – 99% chance that the differences between the groups of calculated RIA values are not due to random sampling. The significance codes provided in the R outputs of the ANOVA analysis are 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’, and so the ‘.’ symbol is shown on the boxplots in Figure 4.32.

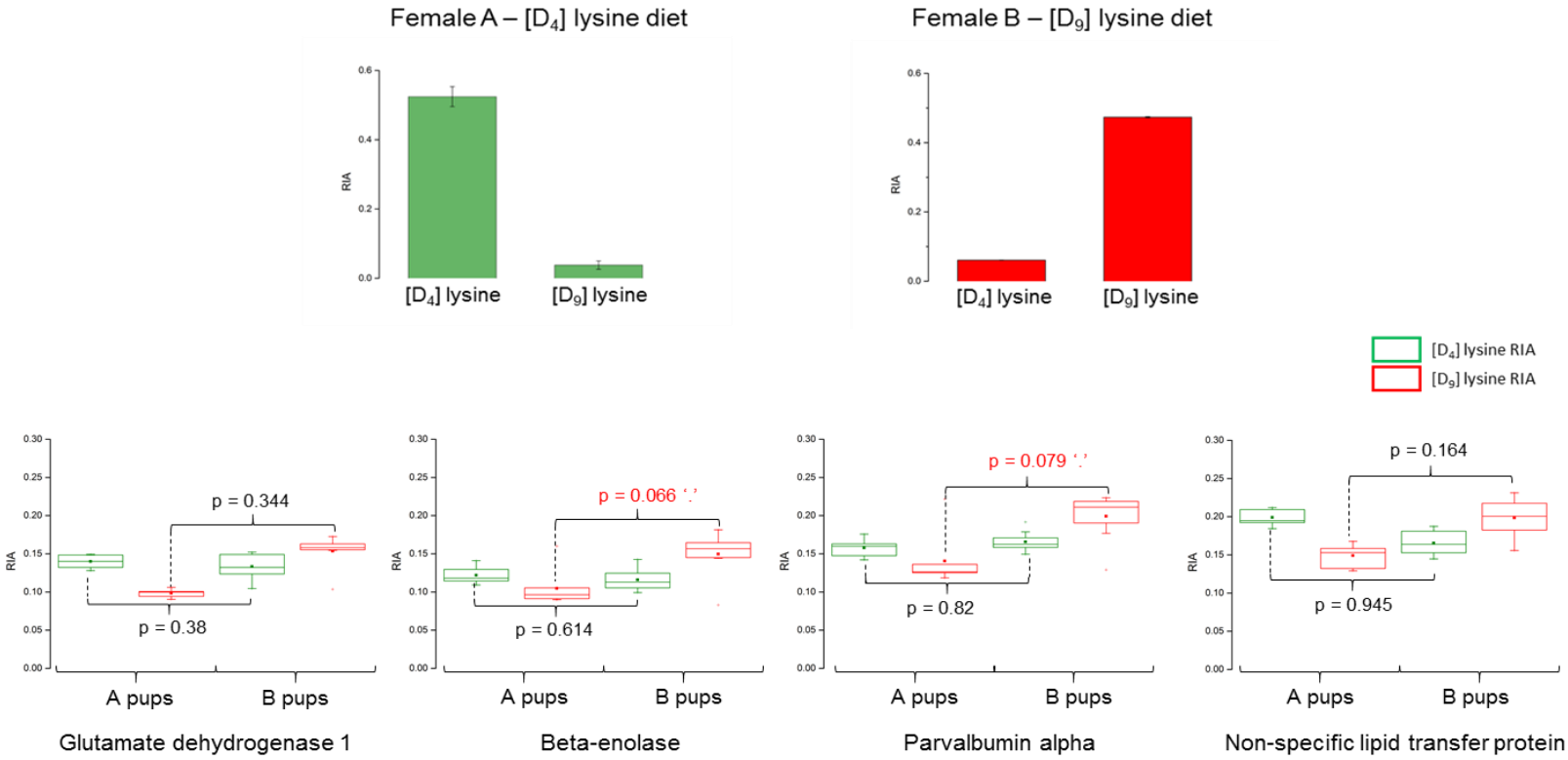


Figure 4.34 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 3 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 3), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs. Those highlighted in red, followed by the '.' symbol, indicate a significant difference in RIA values between the two litters.

Litter 4 (from pair 4)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.35). A and B litter sizes were similar (2 pups and 3 pups, respectively), with litter A born 3 days prior to litter B.

The [D_4] lysine RIAs for each protein are higher than the [D_9] lysine RIAs, indicating that female A consistently invests significantly more than female B in the communal litter (Welch two sample t-test: t = 3.2676, df = 37.998, p = 0.0023). There was no significant difference between the [D_4] lysine precursor RIAs in A and B pups (Welch two sample t-test – glutamate dehydrogenase 1: P = 0.16; beta-enolase: P = 0.335; parvalbumin alpha: P = 0.601; non-specific lipid transfer protein: P = 0.129). There is also no significant difference between the [D_9] lysine precursor RIAs in A and B pups (Welch two sample t-test – glutamate dehydrogenase 1: P = 0.535; beta-enolase: P = 0.456; parvalbumin alpha: P = 0.442; non-specific lipid transfer protein: P = 0.673). Neither female appeared to invest significantly differently in either litter in the communal nest, but female A appeared to invest significantly more in the communal litter than female B.

Litter 5 (from pair 5)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.36). A and B litter sizes were similar (6 pups and 4 pups, respectively), with litter A born 5 days prior to litter B.

The [D_4] lysine RIAs for each protein are significantly higher than the [D_9] lysine RIAs, with female A consistently investing more than female B in the communal litter (Welch two sample t-test: t = 7.7648, df = 75.227, p = 3.285e-11). There is no significant difference between the [D_4] (ANOVA: P = 0.994) and [D_9] lysine (ANOVA: P = 0.617) precursor RIAs in A and B pups in the low turnover liver protein (glutamate dehydrogenase 1), however in the low turnover muscle protein (beta-enolase), there appeared to be significantly higher [D_4] lysine RIA (from female A) in litter B than her own (llitter A) (ANOVA: P = 0.094), but there is no significant difference between the [D_9] lysine precursor RIAs in A and B pups (ANOVA: P = 0.462). The p-value of 0.094, <0.1 but >0.05, is not deemed highly significant. In the high turnover muscle protein (parvalbumin alpha), there is no significant difference between the [D_4]

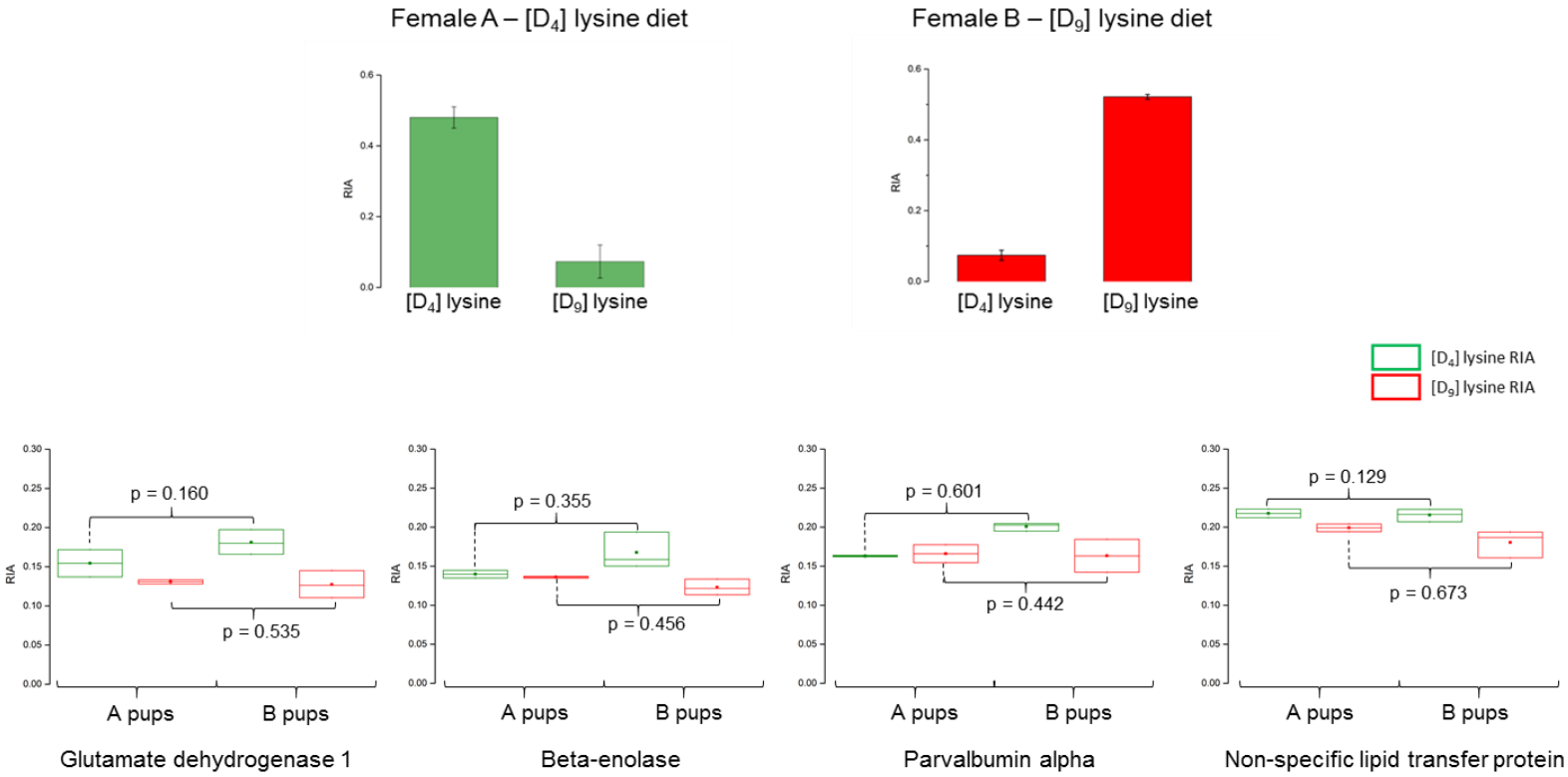


Figure 4.35 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 4 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 4), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

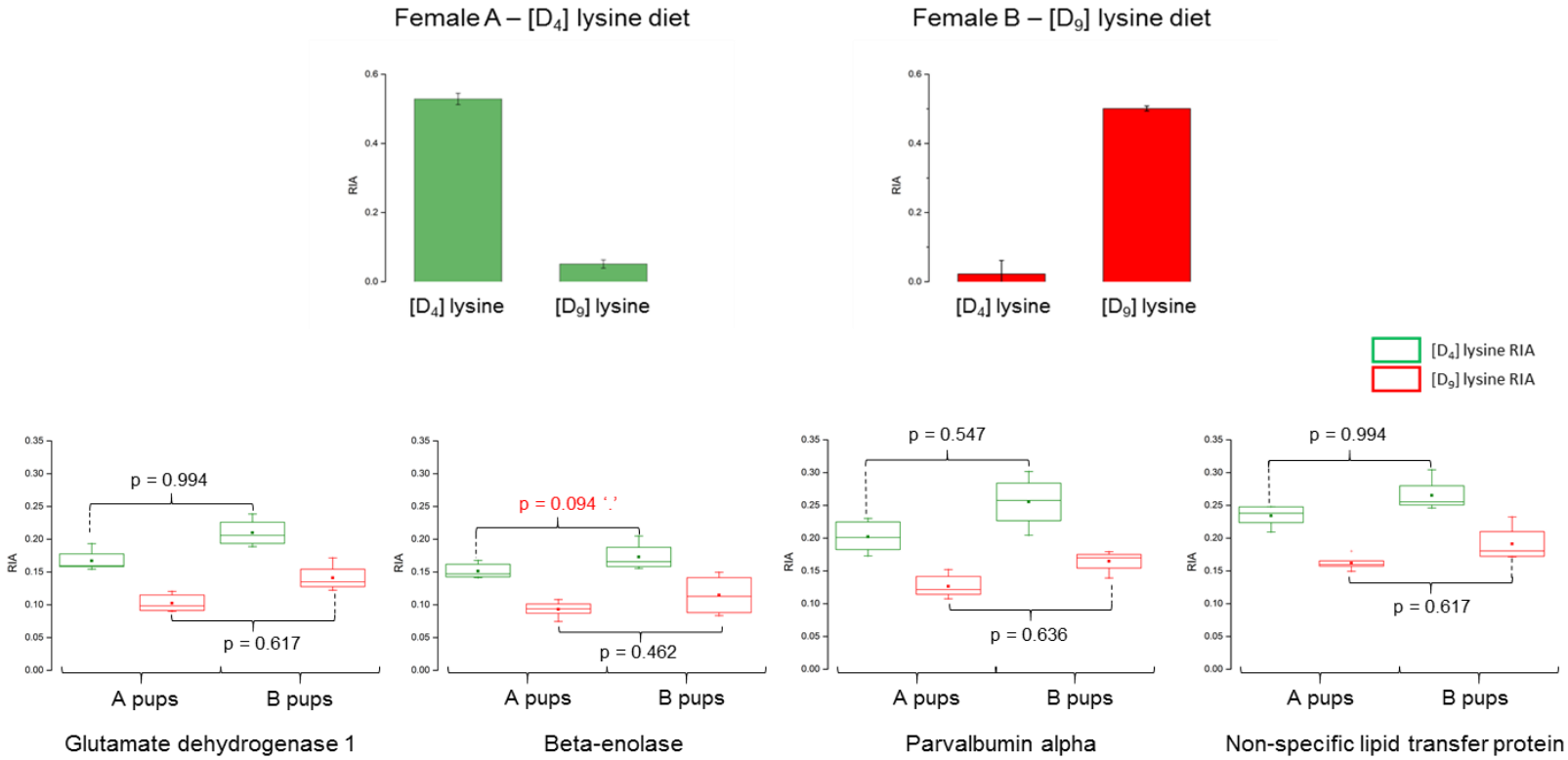


Figure 4.36 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 5 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 5), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs. Those highlighted in red, followed by the '.' symbol, indicate a significant difference in RIA values between the two litters.

(ANOVA: P = 0.547) and [D_9] lysine (ANOVA: P = 0.636) precursor RIAs in A and B pups, and this was the case in the high turnover liver protein (non-specific lipid transfer protein) ([D_4] lysine ANOVA: P = 0.994; [D_9] lysine ANOVA: P = 0.617). In this litter, not only did female A invest significantly more than female B in the entire communal litter, the boxplots in Figure 4.34 indicate that female A invested more (the statistical analysis of the precursor RIAs in beta-enolase suggest somewhat significantly) in female B's litters than her own.

Litter 6 (from pair 6)

The precursor RIAs calculated for each protein of interest for pups from female A and for pups from female B were visualised using boxplots (Figure 4.37). A and B litter sizes were similar (5 pups and 6 pups, respectively), with litter A born 4 days prior to litter B.

The [D_9] lysine RIAs for each protein are significantly higher than the [D_4] lysine RIAs, indicating female B consistently invests slightly more than female A in the communal litter (Welch two sample t-test: t = -2.1432, df = 81.871, p = 0.035). However, there was no significant difference between the [D_4] lysine precursor RIAs in A and B pups (ANOVA –glutamate dehydrogenase 1: P = 0.862; beta-enolase: P = 0.344; parvalbumin alpha: P=0.633; non-specific lipid transfer protein: P=0.99). There is also no significant difference between the [D_9] lysine precursor RIAs in A and B pups (ANOVA –glutamate dehydrogenase 1: P=0.352; beta-enolase: P=0.166; parvalbumin alpha: P=0.68; non-specific lipid transfer protein: P=0.306). Neither female appeared to invest significantly differently in either litter in the communal nest, but female B invested more in the communal litter than female A.

In three of the six unrelated female pairs, both females appeared to invest similarly in the entire communal litter. In one of the pairs where one female invested more in the entire litter than the other (pair 5 of litter 5), one female invested 'highly' significantly more than the other, and in the other two pairs, one female invested significantly more than the other in the communal litter. Again, there were no obvious reason for these differences. In four of the six unrelated pairs, no female appeared to discriminate between their own pups and the other female's pups in terms of investment. In one pair (pair 3 of litter 3), there appeared to be a significant difference between pups A and B in the [D_9] lysine RIAs calculated in both the high and low turnover muscle proteins. In litter 5, there was a significant difference between pups A and B in the

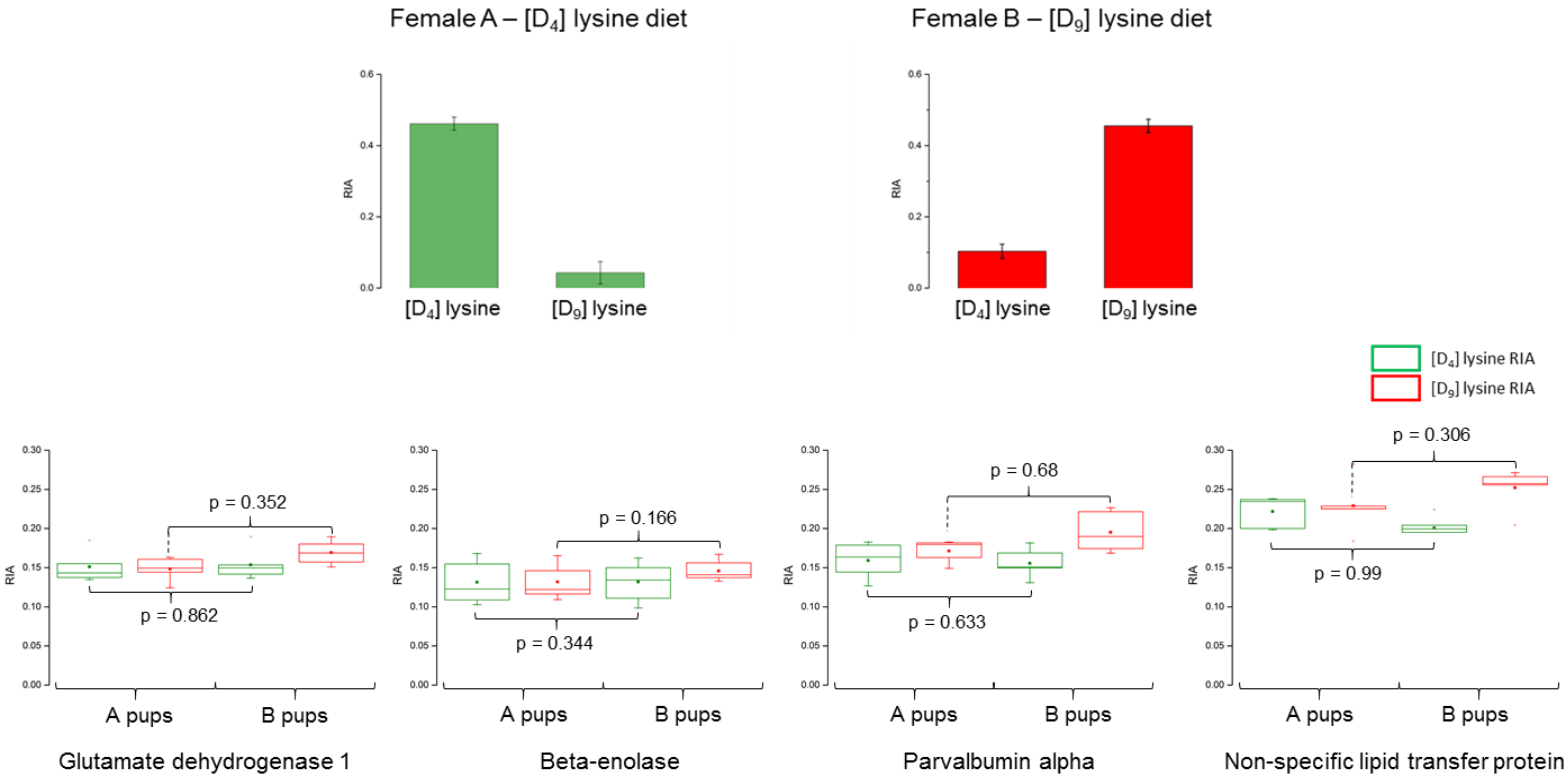


Figure 4.37 The RIAs of heavy amino acid incorporation reached after 7 days in Litter 6 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 6), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter. The p-values are calculated from ANOVA between the [D_4] lysine RIAs calculated for each litter in the nest, for each protein, and the same between the [D_9] lysine RIAs.

[D_4] lysine RIAs calculated in just one protein (the low turnover muscle protein beta-enolase). Since these communal litters were large (a total of 15 pups in communal litter 3 and 10 in litter 5), the differences were not consistent amongst the four proteins, and the differences were not highly significant, it is likely that the differences are due to random sampling and not discrimination in investment by either female.

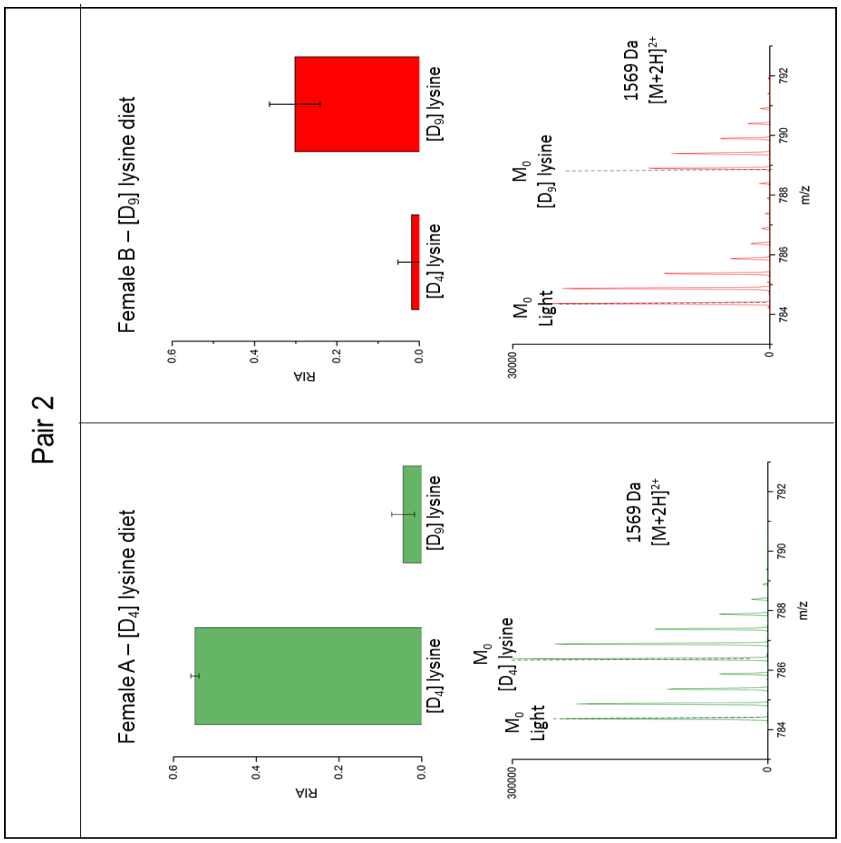
4.2.4.4 Communal nursing of single litters

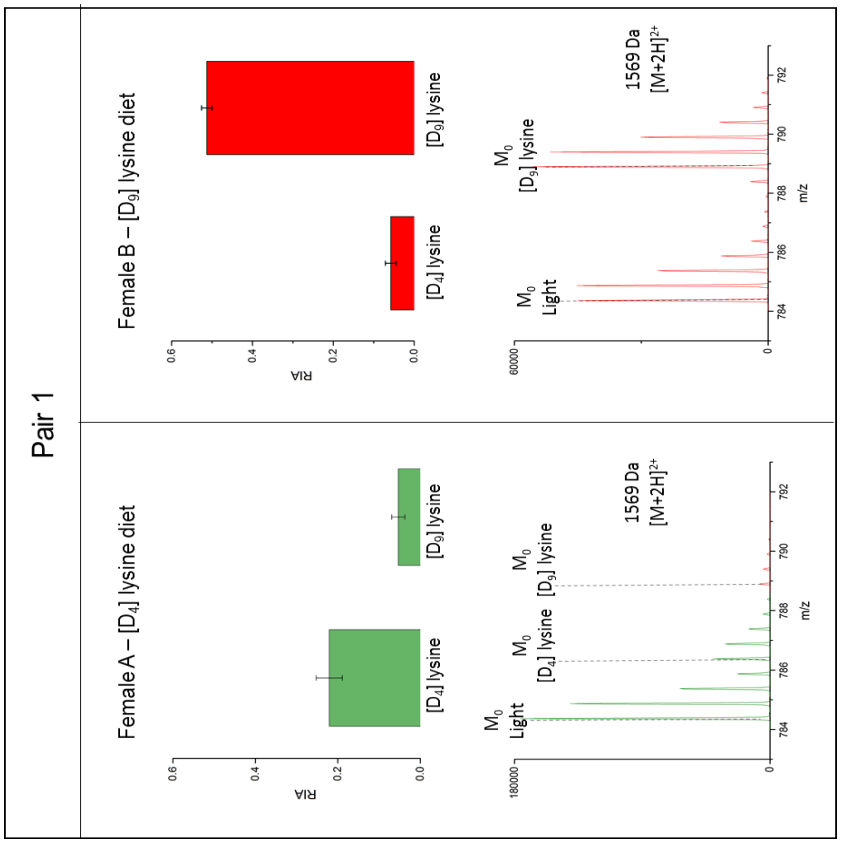
As stated in section 4.2.4.1, a total of 8 related female pairs and 9 unrelated female pairs were put together for breeding, but only 7 related and 6 unrelated female pairs gave birth to and raised communal litters. In the other 4 pairs, only pups from one female survived to the beginning of the experiment (day 1 of females being fed a labelled diet). Since both females in each of these four pairs had given birth, all were lactating, and so it was of interest whether the females with no surviving pups invested in the pups in the communal nest who all belonged to the other female. One pair of females were related, and three pairs were unrelated.

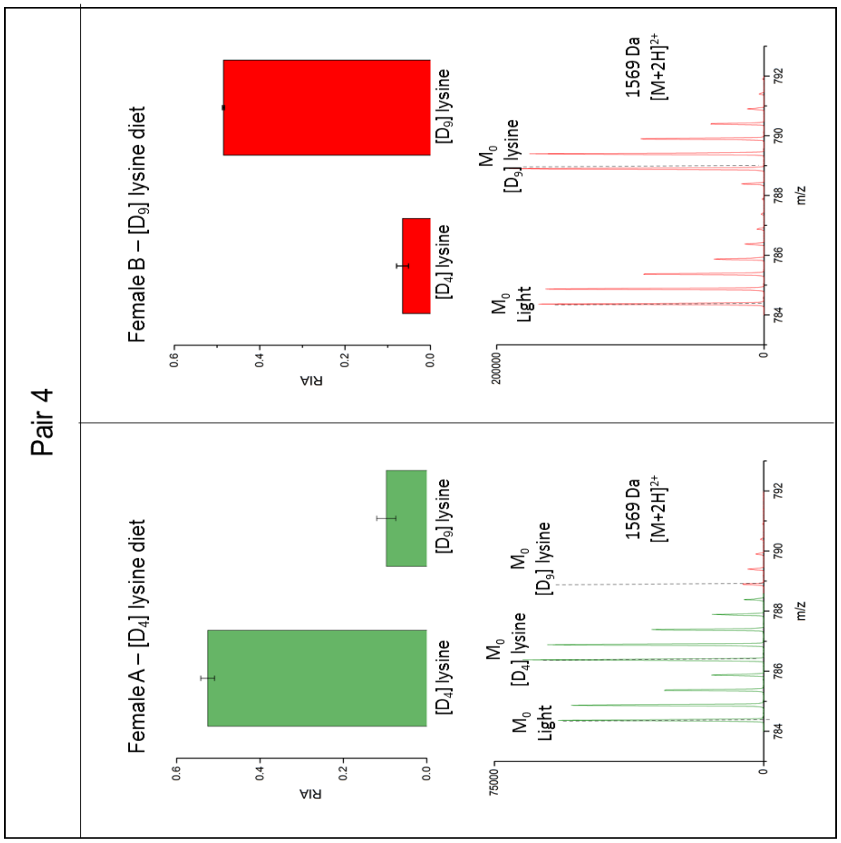
Protein in the mothers' urine was digested in-solution with Lys-C and analysed using LC-MS, and all analysis and calculations were the same as described in section 4.2.4.2. In pair 1 (related females), the female assigned the [D_4] lysine diet (female A) appeared to have a partially [D_4] lysine labelled precursor pool, whilst female B appeared to have a fully [D_9] lysine labelled precursor pool. Female A appeared to consume some of the unassigned [D_9] lysine diet (Figure 4.36). In the three unrelated female pairs, all females assigned the [D_4] lysine labelled diet (female A) appeared to have a fully [D_4] lysine labelled precursor pool. Female A of pair 4 appeared to consume some of the unassigned [D_9] lysine diet (Figure 4.38). Female B in pairs 2 and 3 appeared to have a partially [D_9] lysine labelled precursor pool, and female B in pair 4 appeared to have a fully [D_9] lysine labelled precursor pool. Out of these three females, only female B in pair 3 appeared to consume some of the [D_4] lysine diet (Figure 4.38). After analysis of the pup samples, it was apparent that the incomplete labelling of some females, and the fact that three females appeared to consume some of the unassigned diet, did not need to be taken into consideration, since only the mother of the surviving pups invested in the litter in each case, with evidence of only her label being incorporated into pup tissues (Figure 4.39).

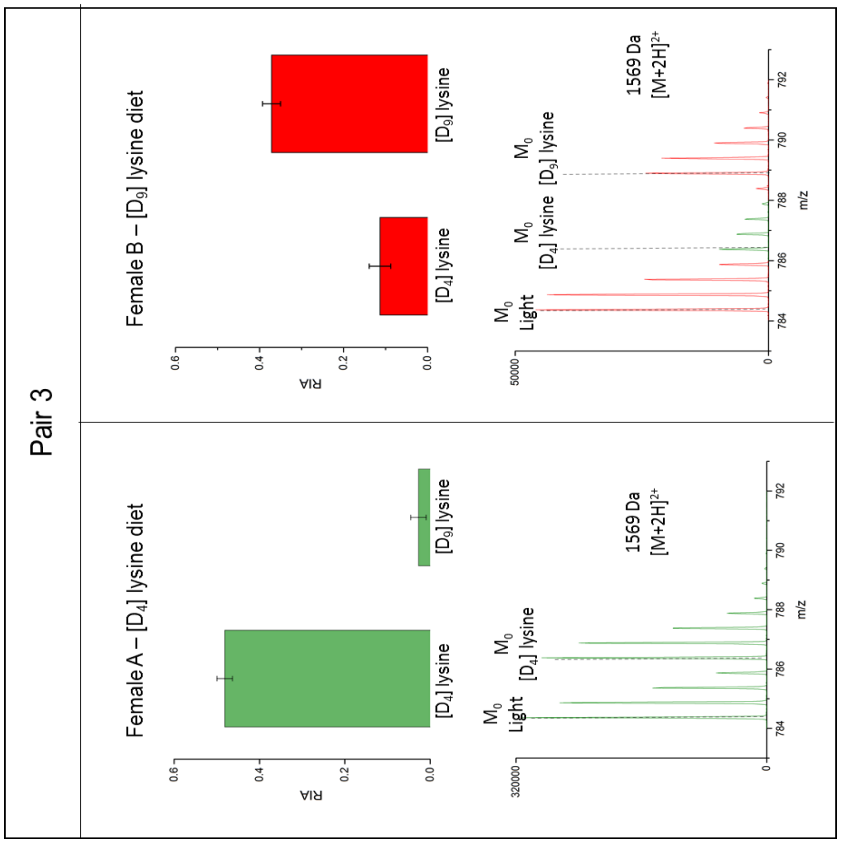
From each pup in each of the four non-communal litters from pairs 1 - 4, after the 7-day labelling experiment, liver samples were taken, homogenised, digested with Lys-C and analysed using LC-MS. Muscle samples were taken from pups in litters 2, 3

Figure 4.38 The RIA of heavy amino acid incorporation reached after 7 days in pairs of female mice assigned either [D_4] lysine or [D_9] lysine labelled diet in nests where only one litter survived to the experiment (pages 224 – 225).
Urine samples were taken from the BALB/c females on the final day of the communal nursing experiment and were subjected to in-solution digestion with Lys-C and analysed using LC-MS. The RIA was calculated from the most abundant MUP peptide in each urine sample. The RIAs for [D_4] lysine and [D_9] lysine were calculated for three MUP peptide in each urine sample (error ± SD, n=3). The isotope distribution patterns for the 1569 Da $[M+2H]^{2+}$ MUP peptide are shown for each female urine sample – the [D_4] lysine peptides are shown in green and the [D_9] lysine peptides are shown in red. The unlabelled peptides are shown in green for females assigned the [D_4] lysine diet, and red for females assigned the [D_9] lysine labelled diet.


Pair 2
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da [M+2H]2+
m/z


Pair 1
Female B – [D9] lysine diet
Female A – [D4] lysine diet
RIA
[D4] lysine
[D9] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da [M+2H]2+
m/z


Pair 4
Female B – [D9] lysine diet
RIA
[D9] lysine
[D4] lysine
M0 [D9] lysine
M0 Light
1569 Da [M+2H]2+
m/z
Female A – [D4] lysine diet
M0 [D4] lysine
M0 [D9] lysine


Pair 3
Female B – [D9] lysine diet
RIA
[D9] lysine
[D4] lysine
M0 Light
M0 [D4] lysine
M0 [D9] lysine
1569 Da [M+2H]2+
m/z
Female A – [D4] lysine diet

and 4, and were analysed in the same way. All analysis was carried out as described in section 4.2.4.1.

Non-communal litter 1 (from related pair 1)

The 7 pups in the non-communal litter all belonged to female B. The precursor RIAs calculated for each liver protein of interest for pups were visualised using boxplots (Figure 4.40). Only female B invested in the litter – the [D_9] lysine RIAs in the two selected liver proteins in pups averaged at 0.13 in the low turnover protein (glutamate dehydrogenase 1) and at 0.20 in the high turnover protein (non-specific lipid transfer protein). Female A did not invest in any of her sister's pups, with no evidence of [D_4] lysine in the two liver proteins. The Welch two sample t-test confirmed that this difference in investment from the two females was highly significant (t = -18.4109, df = 13, p = 1.077e-10).

Non-communal litter 2 (from unrelated pair 2)

In this unrelated pair, all 4 pups in the non-communal litter belonged to female A. The precursor RIAs calculated for each liver and muscle protein of interest for pups were visualised using boxplots (Figure 4.41). Only female A invested in the litter – female B invested in none of Female A's pups, with no evidence of [D_9] lysine in any of the proteins. The [D_4] lysine RIAs in the pups averaged at 0.25, ranging from approximately 0.21 in the low turnover proteins (glutamate dehydrogenase 1 and beta-enolase) to 0.32 in the high turnover liver protein (non-specific lipid transfer protein). As expected, the Welch two sample t-test confirmed that this difference in investment from the two females was highly significant (t = 21.8468, df = 15, p = 8.743e-13).

Non-communal litter 3 (from unrelated pair 3)

The 2 pups in the non-communal litter belonged to female A. The precursor RIAs calculated for each liver and muscle protein of interest for pups were visualised using boxplots (Figure 4.42). Again, only female A invested in the litter – female B invested in none of Female A's pups, with no evidence of [D_9] lysine in any of the proteins. The [D_4] lysine RIAs in the pups averaged at 0.26, ranging from approximately 0.21 in the low turnover proteins (glutamate dehydrogenase 1 and beta-enolase) to 0.32 in the high turnover liver protein (non-specific lipid transfer protein). The Welch two

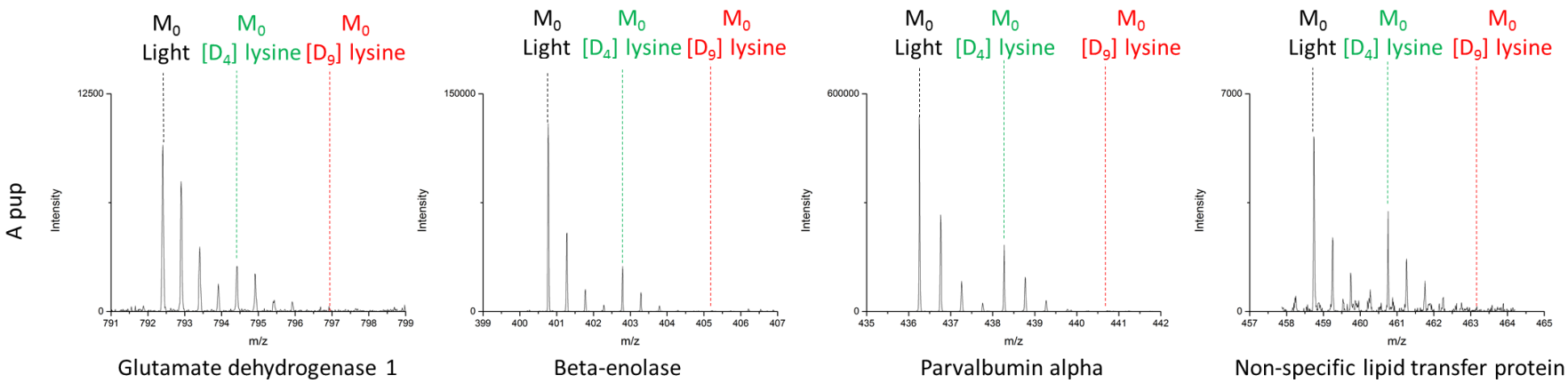


Figure 4.39 Examples of the heavy amino acid incorporation reached after 7 days in the high and low turnover liver and muscle proteins in pups from a non-communal litter, as a result of investment from their mother only, who was assigned the [D_4] lysine labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. Shown are the unlabelled ('light') and labelled ('heavy') peptide profiles of a glutamate dehydrogenase 1 peptide (low turnover in liver), a beta-enolase peptide (low turnover in muscle), a parvalbumin alpha peptide (high turnover in muscle), and a non-specific lipid transfer protein peptide (high turnover in liver), confirming the incorporation of only [D_4] lysine (profiles labelled green) and not [D_9] lysine (red) into pup tissue samples as a result of investment from their own mother only (female A). This figure shows the spectra of tissue samples taken from an 'A pup' (whose mother is female A) from Litter 2.

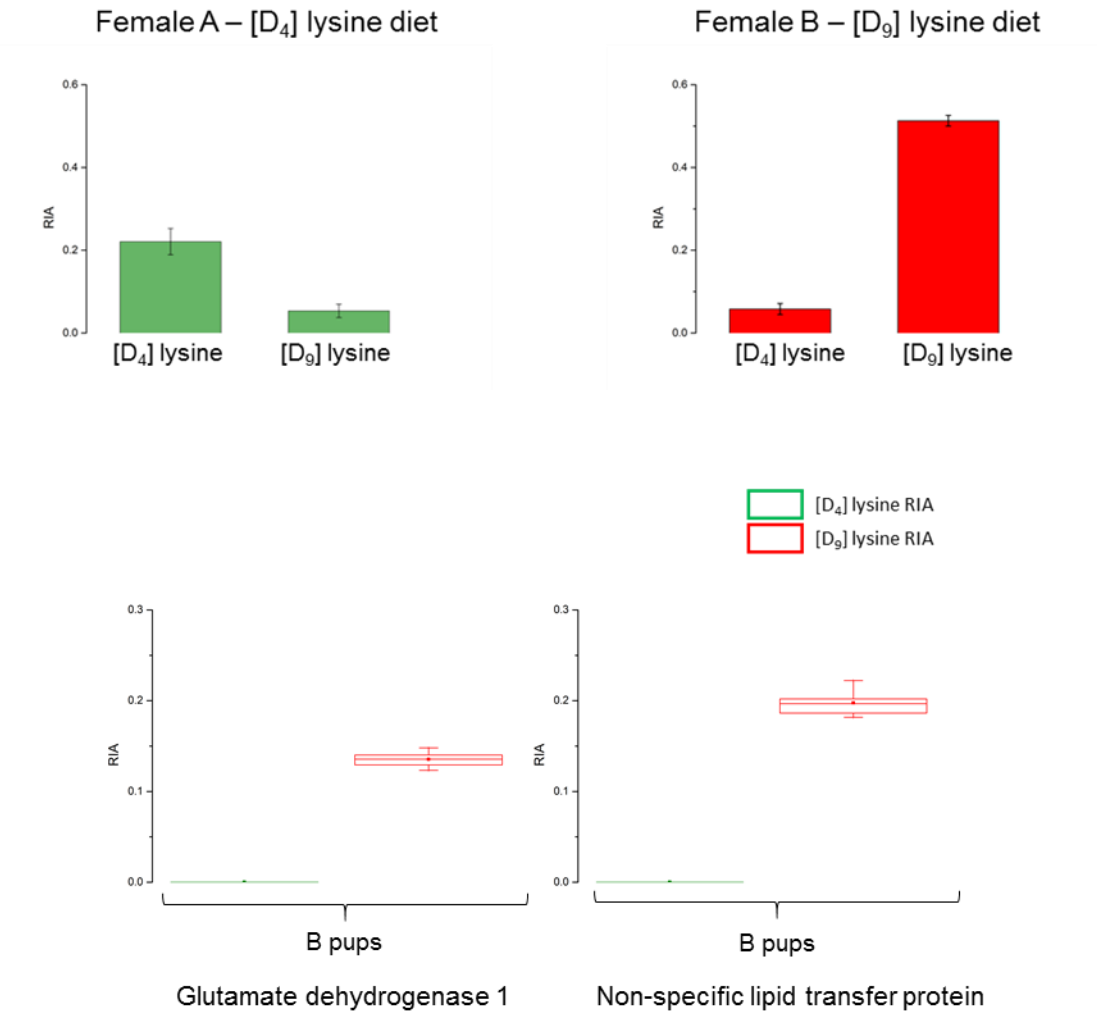


Figure 4.40 The RIAs of heavy amino acid incorporation reached after 7 days in non-communal Litter 1 pup liver samples as a result of investment from related mothers (Pair 1), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter.

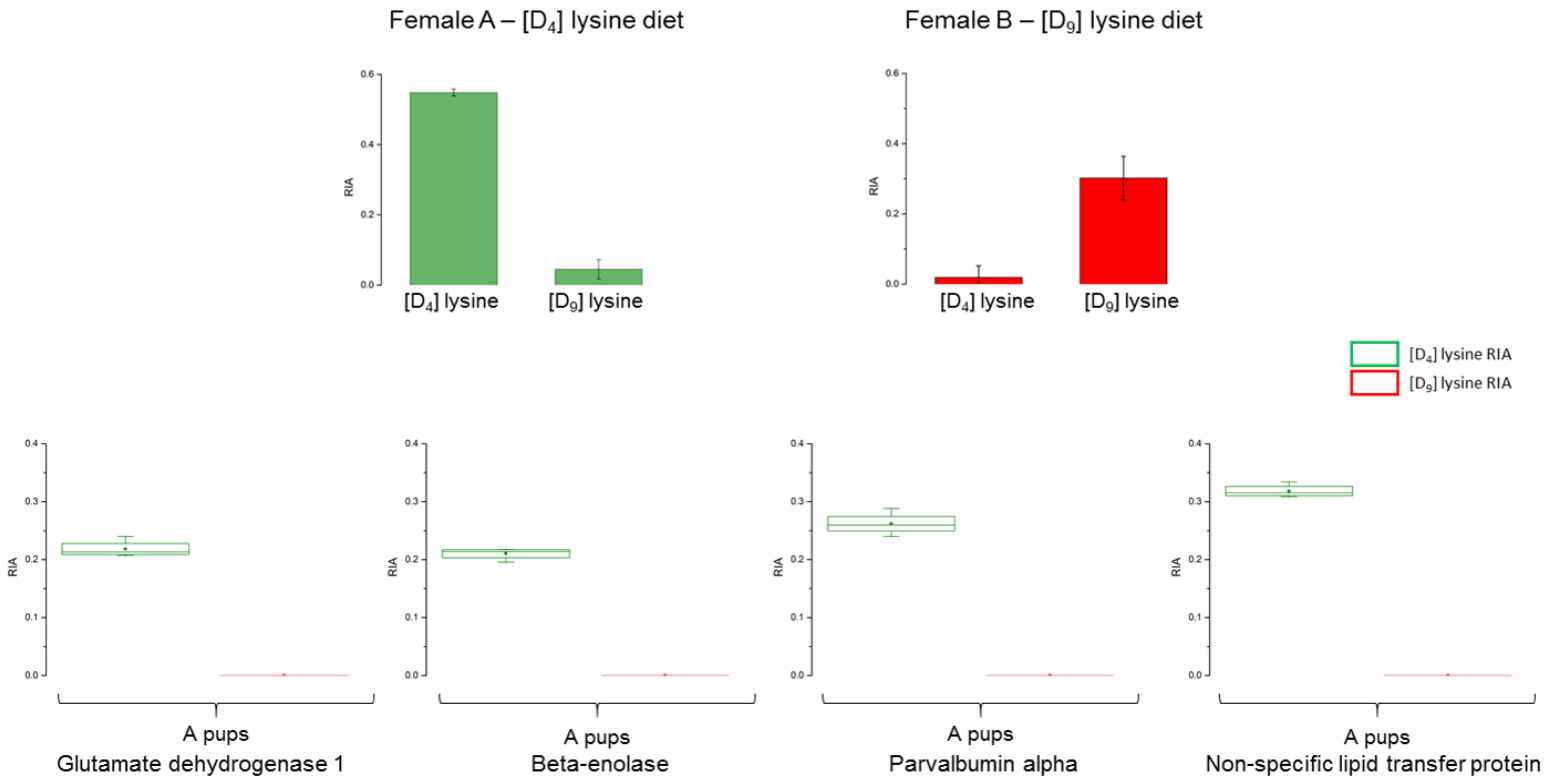


Figure 4.41 The RIAs of heavy amino acid incorporation reached after 7 days in non-communal Litter 2 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 2), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter.

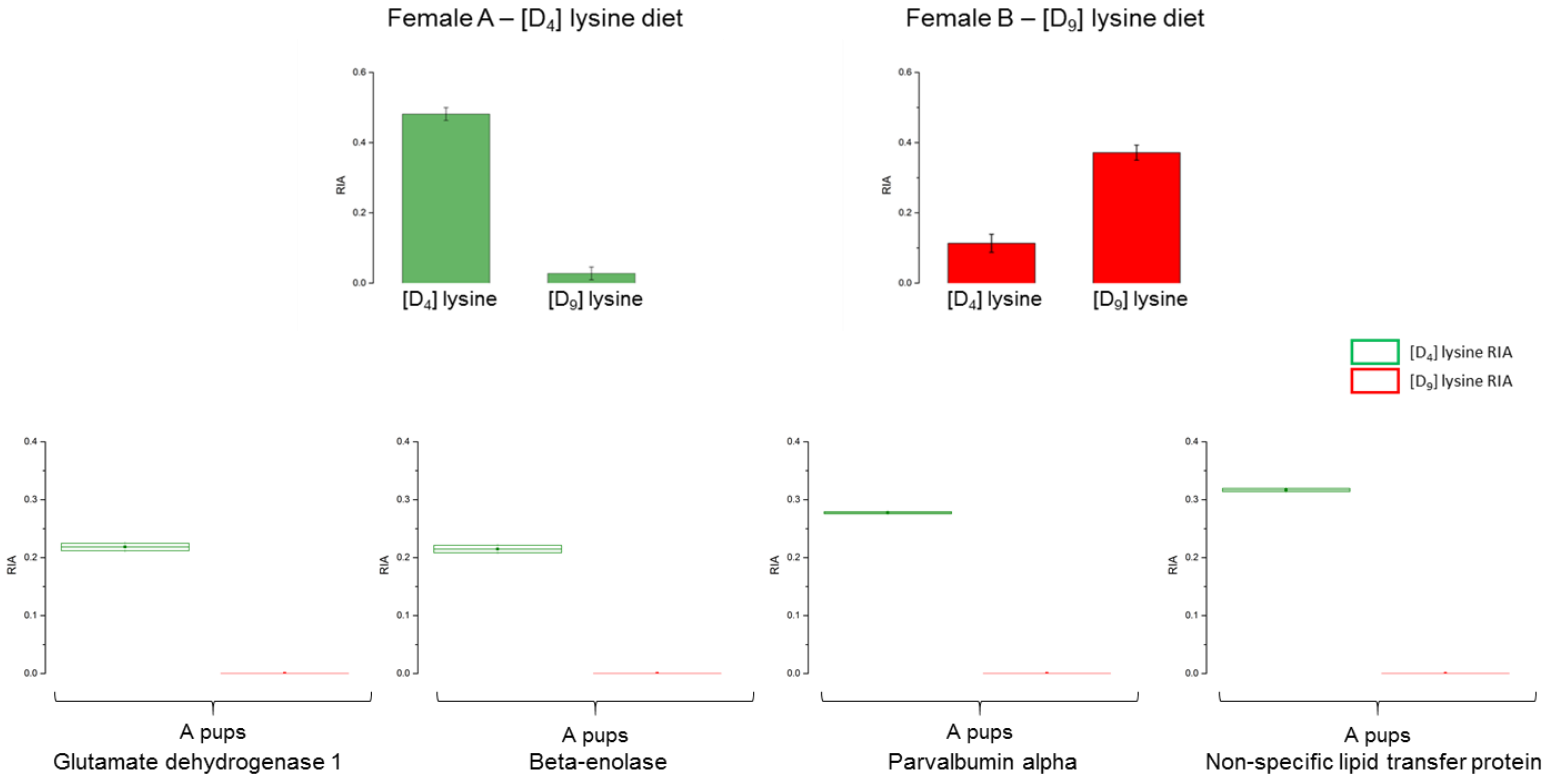


Figure 4.42 The RIAs of heavy amino acid incorporation reached after 7 days in non-communal Litter 3 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 3), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter.

sample t-test confirmed that this difference in investment from the two females was highly significant (t = 15.8506, df = 7, p = 9.645e-07).

Non-communal litter 4 (from unrelated pair 4)

The 2 pups in the non-communal litter belonged to female A. The precursor RIAs calculated for each liver and muscle protein of interest for pups were visualised using boxplots (Figure 4.43). As with the other three non-communal litters, only the mother of the pups invested in the litter – female B invested in none of Female A's pups, with no evidence of [D_9] lysine in any of the proteins. The [D_4] lysine RIAs in the pups averaged at 0.27, ranging from approximately 0.21 in the low liver turnover protein (glutamate dehydrogenase 1) to 0.32 in the high turnover liver protein (non-specific lipid transfer protein). The Welch two sample t-test confirmed that this difference in investment from the two females was highly significant (t = 21.7305, df = 11, p = 2.19e-10).

In the related pair and the three unrelated female pairs, only the mother of the single litter of pups appeared to invest in the litter. As a result, the differences in female investment in all four non-communal litters were deemed highly significant. As all the investing females had fully, correctly labelled precursor pools, the fact that the non-investing females did not have fully, correctly labelled precursor pools did not need to be taken into account when calculating precursor RIA values in pup tissue samples.

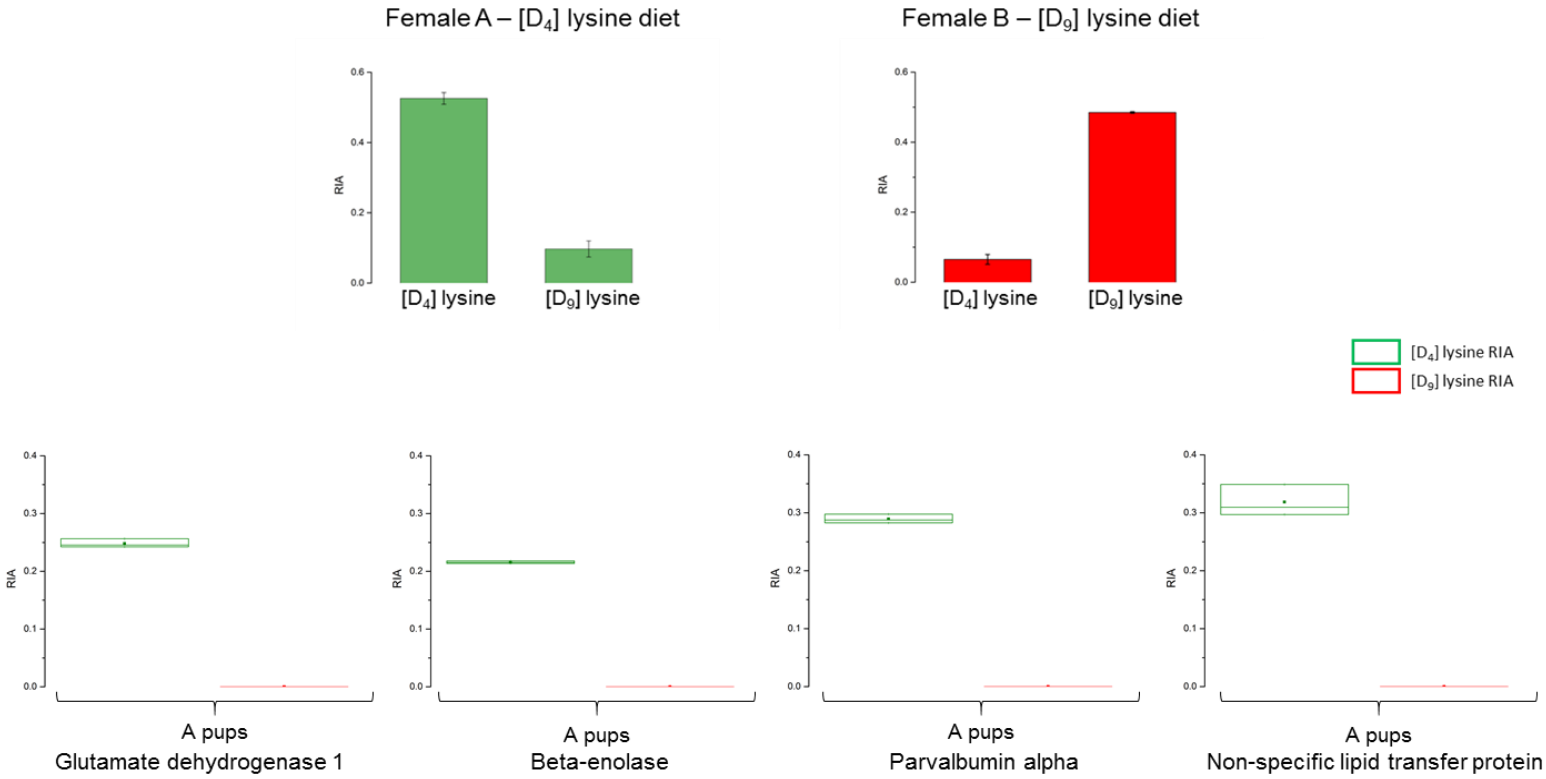


Figure 4.43 The RIAs of heavy amino acid incorporation reached after 7 days in non-communal Litter 4 pup liver and pup muscle samples as a result of investment from unrelated mothers (Pair 4), who were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet.

Tissue samples were taken from all pups on the final day of the communal nursing experiment and were subjected to homogenisation and in-solution digestion with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the average RIAs of the each litter in the communal nest, along with the variation of calculated RIAs amongst the litter.

4.3 Conclusions

The aim of this study was to determine the proportion of lactation investment from communally nursing house mice to determine any discriminatory factors, mainly the relatedness of the female pairs (mothers) in the communal nest. This was done by successfully incorporating two different labelled amino acids, one into each female in the pair, via a semi-synthetic diet until whole animal metabolic labelling had been achieved. The labelled amino acid was then transferred to the pups which the females invested in via their milk, and so the amount of each label present in pup tissues determined the proportion of investment from each of the two females. Label incorporation into proteins in the mother's liver, milk and pup's tissues was confirmed by mass spectrometric analysis of female urine, pup stomach contents and pup tissue samples. The rate of incorporation of label into females was studied, followed by the development of a feeding mechanism that allowed two females living together to consume two differently labelled diets. The incorporation of label into proteins in female's milk was confirmed, as was the incorporation into proteins in pup tissues. These preparatory experiments allowed the successful tracking of labels, and therefore investment, from pairs of females in the pups in their communal litter. Statistical analysis enabled the determination of any significant differences in investment from either mother in the litters present in the communal nest.

Incorporating labelled amino acids into adult female BALB/c mice via a semi-synthetic diet was successful – diets were 50% labelled with the chosen synthetic amino acid, so that mice found the diet palatable and so whole animal metabolic labelling was achieved. To determine the rate of full incorporation of the labelled amino acids into all proteins in the animal, female mice were fed either a [$^{2}H_{8}$] valine or [$^{13}C_{6}$] lysine labelled diet. Urine samples were taken from them daily, and mass spectrometric analysis of MUPs in the urine allowed analysis of the liver precursor pool. Labelled and unlabelled MUP peptides were identified in the mass spectra, and precursor RIA values for the protein were calculated from the signal intensities of the labelled and unlabelled peptides. A precursor RIA of 0.5 indicated a fully labelled precursor pool (as the diet provided was 50% labelled). Fully labelled precursor pools were reached in the animals within 2 – 3 days of being fed a labelled diet: at this point, all proteins being synthesised in the liver and all proteins degrading contain both labelled and unlabelled valine/lysine, returning an almost equal amount of these to the precursor pool (Figure 4.44). In the feeders analysis and milk labelling experiments mice were fully labelled within 1 – 2 days of being fed a labelled diet.

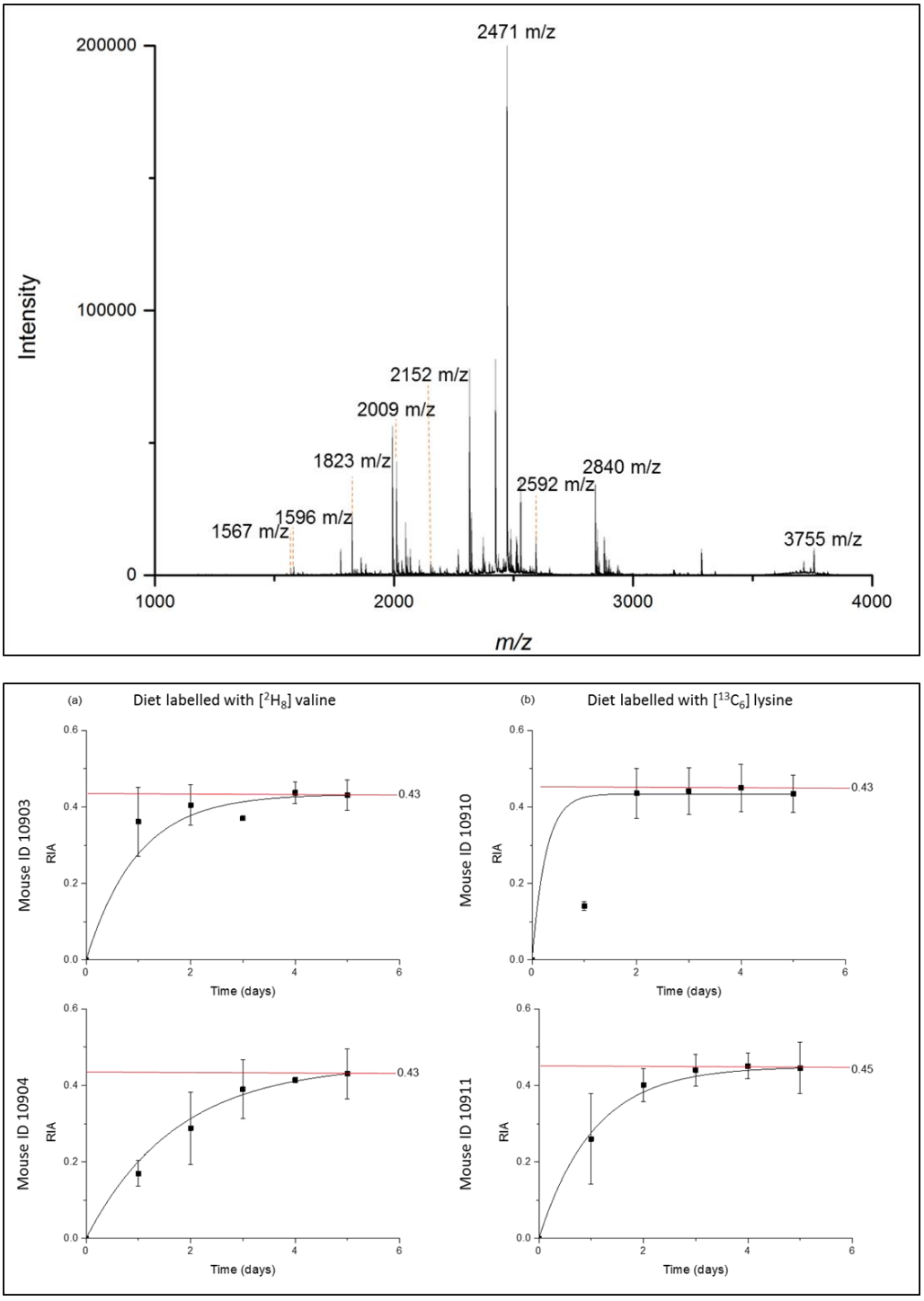


Figure 4.44 Summary of incorporation of labelled amino acids into adult female BALB/c mice via labelled diets, using copies of Figure 4.2 and Figure 4.6.
Top panel: Peptide mass fingerprint of female BALB/c mouse urine prior to labelled diet. Whole urine was digested with Lys-C and subjected to MALDI-ToF-MS analysis. Labelled are the MUP peptides identified by database search via Mascot search engine. Bottom panel: Urine samples were taken from the four BALB/c females at various time intervals over a six-day period of being fed a labelled diet. Samples were digested with Lys-C and analysed using MALDI-ToF-MS. An average RIA was calculated from the three most abundant valine-containing (a) and lysine-containing (b) peptides for each day (error ± SD). Average maximum RIAs reached for each mouse are shown by a red line.

A feeding mechanism in the communal nesting cages was developed at Leahurst which allowed two females, living in the same cage, to access two differently labelled diets whilst not having access to the other diet. The feeding mechanism was tested with three pairs of mice – in each pair, one female was assigned a [$^{2}H_{8}$] valine diet, and the other, a [$^{13}C_{6}$] lysine diet. Over the five day experiment, urine samples were taken from each animal daily, and the mass spectrometric analysis of MUPs in the urine allowed analysis of the liver precursor pool. [$^{2}H_{8}$] valine and [$^{13}C_{6}$] lysine labelled and unlabelled MUP peptides were searched for in the mass spectra, with all MUP peptides analysed for the presence of the 'correct' label (the label arising from the labelled diet which the mouse was assigned), and also the presence of the 'wrong' label (the label from the diet which the other mouse in the pair was assigned to). Calculated precursor RIA values confirmed that mice were fully labelled with their assigned diet after 1 – 2 days of the experiment, with no evidence of any animal consuming any of the wrong, unassigned diet. The mechanism allowed mice to access their assigned diet, whilst preventing it from accessing any of the unassigned diet (Figure 4.45). Come the communal nursing experiment, with this cage/feeding mechanism set-up, two female mice would be able to live, nest, and communally nurse their litters, allowing their investment in the pups can be tracked via the transfer of two different labels from the mothers to the pups.

The rate at which labelled amino acids were incorporated into the milk of lactating females was assessed by feeding three pairs of communally nesting females a [$^{13}C_{6}$] lysine labelled diet once both females in the pair had given birth to their litter. Analysis of the mothers' urine, which was taken daily during the labelling stage, showed that MUPs were fully labelled within 1 - 2 days of being fed the labelled diet. The stomach contents of a single pup were taken daily for the analysis of the labelling of milk proteins. In the sample taken from a pup immediately prior to the labelling stage, no [$^{13}C_{6}$] lysine labelled peptides were observed in the MALDI-ToF spectrum, confirming that no label was present in the mothers' milk prior to being fed the labelled diet. Over the course of the experiment, the lactation elevation protein present in the milk samples labelled at the same rate as MUPs in the mothers' urine. A slightly lower precursor RIA was calculated in the milk samples recovered on days 2 and 3 of the experiment, and this is likely to be due to the time at which these samples were taken: in the initial kinetics experiment, where the labelling of the female liver precursor pool was assessed, precursor RIAs calculated in the MUP peptides were slightly lower in the urine samples taken in the morning than the ones taken in the afternoon/evening

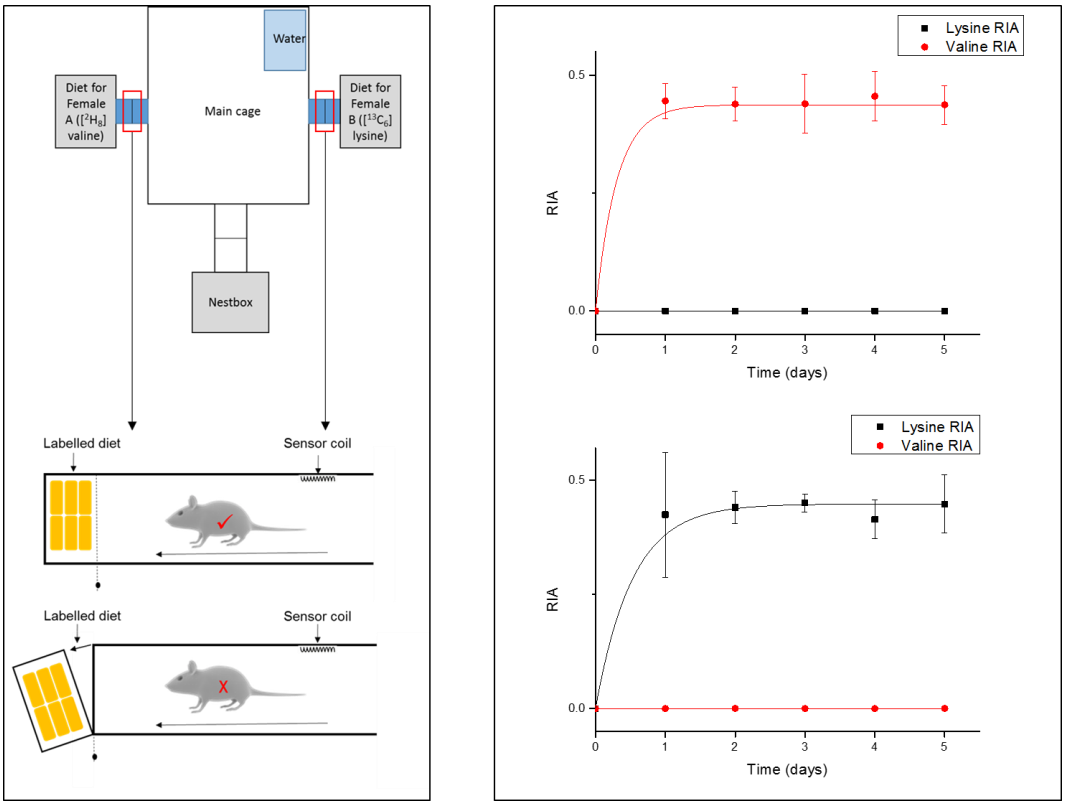


Figure 4.45 Summary of the feeding mechanism developed to allow two communally nesting females to consume differently labelled diets, using copies of Figure 4.7 and Figure 4.9.
Left panel: Two feeding traps, each containing a different diet, are positioned at each end of the cage, with both mice having access to all parts of the cage. Each female mouse has an RFID chip implanted, and this chip can be read by the sensor coil placed at the entrance of each feeding mechanism. If a mouse approaches the correct diet, the feed will remain in reach of the mouse. If a mouse was to approach the incorrect diet, the feeding mechanism would sense this via the chip carried by the mouse, and would subsequently move the diet away from the cage, out of reach of the mouse. Right panel: The rate of heavy amino acid incorporation over 6 days in three mice fed [$^{2}H_{8}$] valine labelled diet (top) and three mice fed [$^{13}C_{6}$] lysine labelled diet (bottom). Three nests with the feeding mechanism contained two female BALB/c mice in each, each mouse in the pair was assigned either the [$^{2}H_{8}$] valine diet or the [$^{13}C_{6}$] lysine diet. Urine samples were taken from the six mice every day over the six-day period, digested with Lys-C and analysed using MALDI-ToF-MS. An average daily RIA was calculated from the two most abundant valine and lysine-containing MUP peptides for each mouse, each day (error ± SD).

(Figure 4.46). This is because mothers will have been inactive during their 'day' period overnight, and so will have been consuming less labelled diet prior to the morning samples being taken. In this experiment, pup stomach contents were taken at 9 am, and so the slightly lower RIAs calculated for the milk proteins are likely to be as a result of the inactivity of the mothers during the previous night. Taking into account that mothers' urine was sampled at 5 pm daily, after their active period, the proteins in the mothers' milk (in the pup stomachs) reaches a similar labelling maximum, at a similar rate, to the protein in the mothers' urine.

The rate at which labelled amino acids were incorporated into the pups was assessed by analysing pup urine, liver and muscle samples, which were taken at the same time as the stomach contents. Precursor RIA values were calculated for two proteins in the urine, and were much lower than those calculated in the pup stomach contents, not reaching a maximum labelling plateau at any point during the experiment. This is because label incorporation into the pups is much slower, as the label is introduced into the pups indirectly, and the rate of growth of pups means that the label is being diluted by a greater amount of unlabelled protein present in the precursor pool. The precursor RIAs in serum albumin in the urine increase throughout the experiment, confirming that the label is being successfully incorporated into the pups. Analysis of the proteins in liver and muscle samples was by LC-MS, whilst all previous samples had been analysed using MALDI-ToF-MS since for protein discovery, the separation of peptides and the use of ESI-QToF allowed the generation of better product ion information and therefore more accurate peptide identification. The liver and muscle samples taken on the final day of the milk labelling experiment were analysed first, to determine the maximum RIAs reached in the top 20 proteins in each sample type. Maximum precursor RIAs in the pup liver ranged from 0.15 to 0.47, and in pup muscle from 0.15 to 0.36 (Figure 4.47). The aim was to find a high and low turnover protein from each tissue type, with high turnover proteins incorporating the label more quickly (high precursor RIAs) than low turnover proteins (low precursor RIA). Proteins with different turnover rates in pup tissues were hoped to give information regarding investment from mothers at different time points during communal nursing. High turnover proteins degrade and return to the protein precursor pool quickly, with newly synthesised labelled proteins taking their place. This high turnover of proteins causes rapid labelling of the precursor pool, so recent changes in label investment can be identified through the analysis of these proteins. The opposite is true for low turnover proteins, meaning that these proteins may be able to report on variations in label

Day of Experiment	Samples Taken		
	Communal Litter 1	Communal Litter 2	Communal Litter 3
0 – Unlabelled diet	Female urine + 1 pup	Female urine	Female urine
1 – [$^{13}C_6$] lysine labelled diet	Female urine	Female urine + 1 pup	Female urine + 1 pup
2 – [$^{13}C_6$] lysine labelled diet	Female urine + 1 pup	Female urine + 1 pup	Female urine
3 – [$^{13}C_6$] lysine labelled diet	Female urine + 1 pup	Female urine	Female urine + 1 pup
4 – [$^{13}C_6$] lysine labelled diet	Female urine	Female urine + 1 pup	Female urine + 1 pup
5 – [$^{13}C_6$] lysine labelled diet	Female urine+ 1 pup	Female urine + 1 pup	Female urine
6 – [$^{13}C_6$] lysine labelled diet	Female urine + 1 pup	Female urine	Female urine + 1 pup

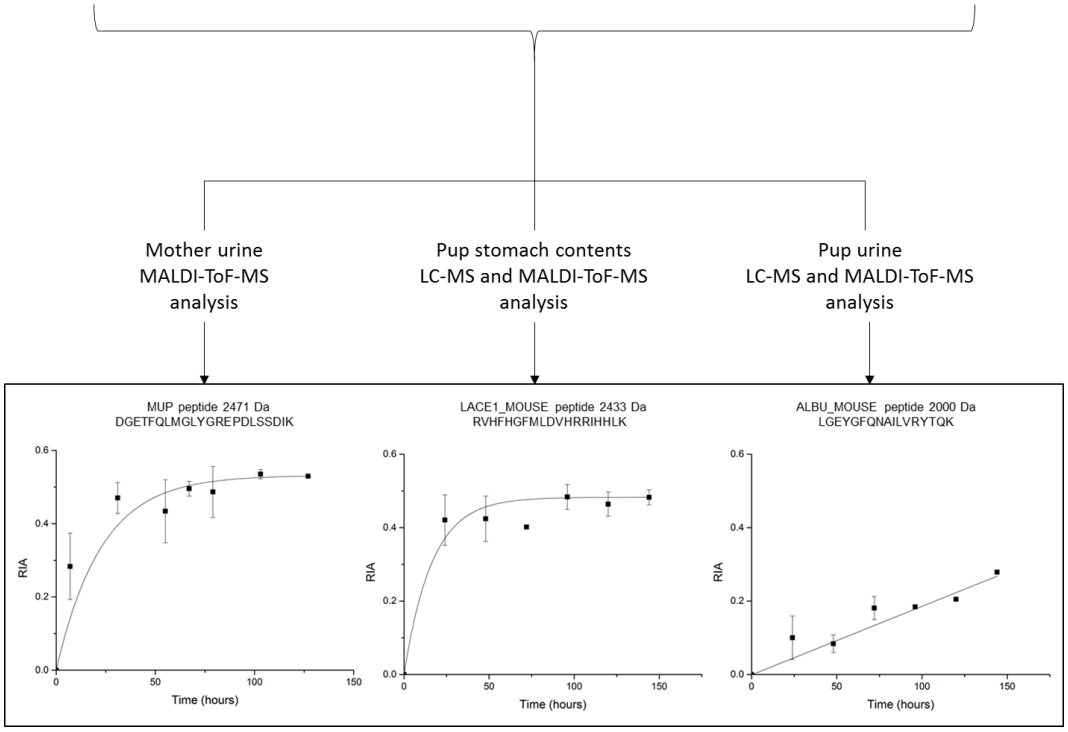


Figure 4.46 Summary of the experimental set up of the milk labelling pilot and the rate of heavy amino acid incorporation over 6 days in adult female urine, pup stomach contents and pup urine.
Three pairs of pregnant females lived and gave birth in in a communal nursing cage, consuming an unlabelled diet. When all females had given birth, a single pup was taken from one of the communal litters for analysis, along with urine samples from all females. All females were then switched to a [$^{13}C_6$] lysine labelled diet. Over days 1 – 6, urine from all females was sampled daily for analysis, and two pups were taken from one of the three communal litters (highlighted in red) for analysis each day over the course of the labelling experiment. Female urine samples were digested with Lys-C and analysed using MALDI-ToF-MS. An average RIA for the three mice calculated from the most abundant lysine-containing peptide for each day (n = 6, error ± SD). Pup stomach contents and urine samples were digested with Lys-C. Samples were subjected to LC-MS analysis for protein discovery, further and analysis and RIA calculations were carried out in the same way as the mothers' urine, using MALDI-ToF-MS.

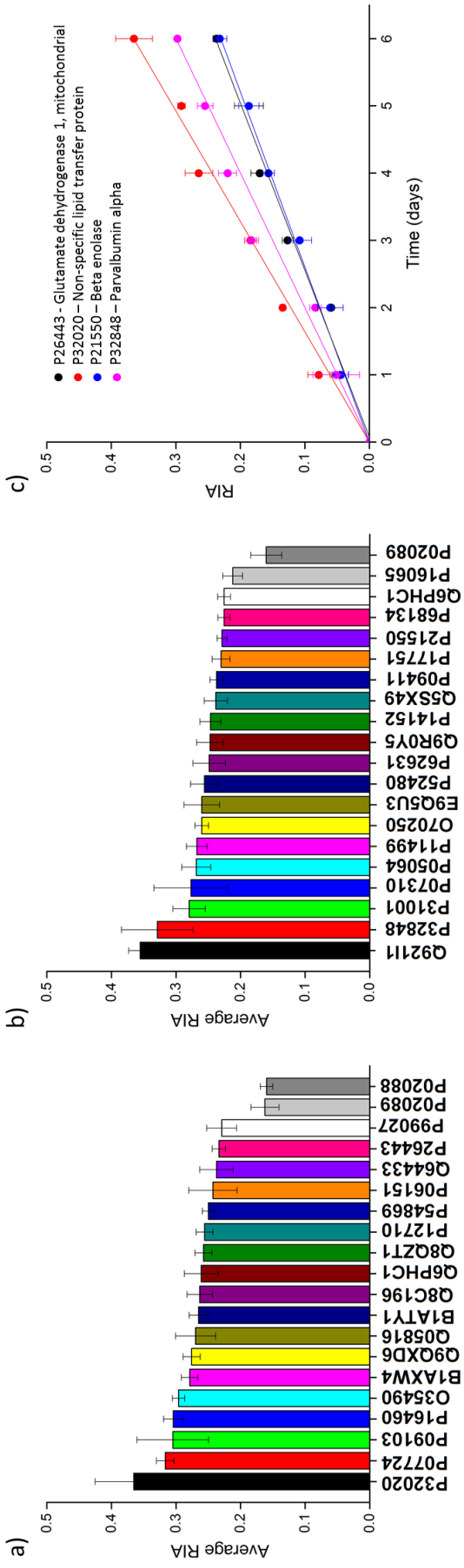


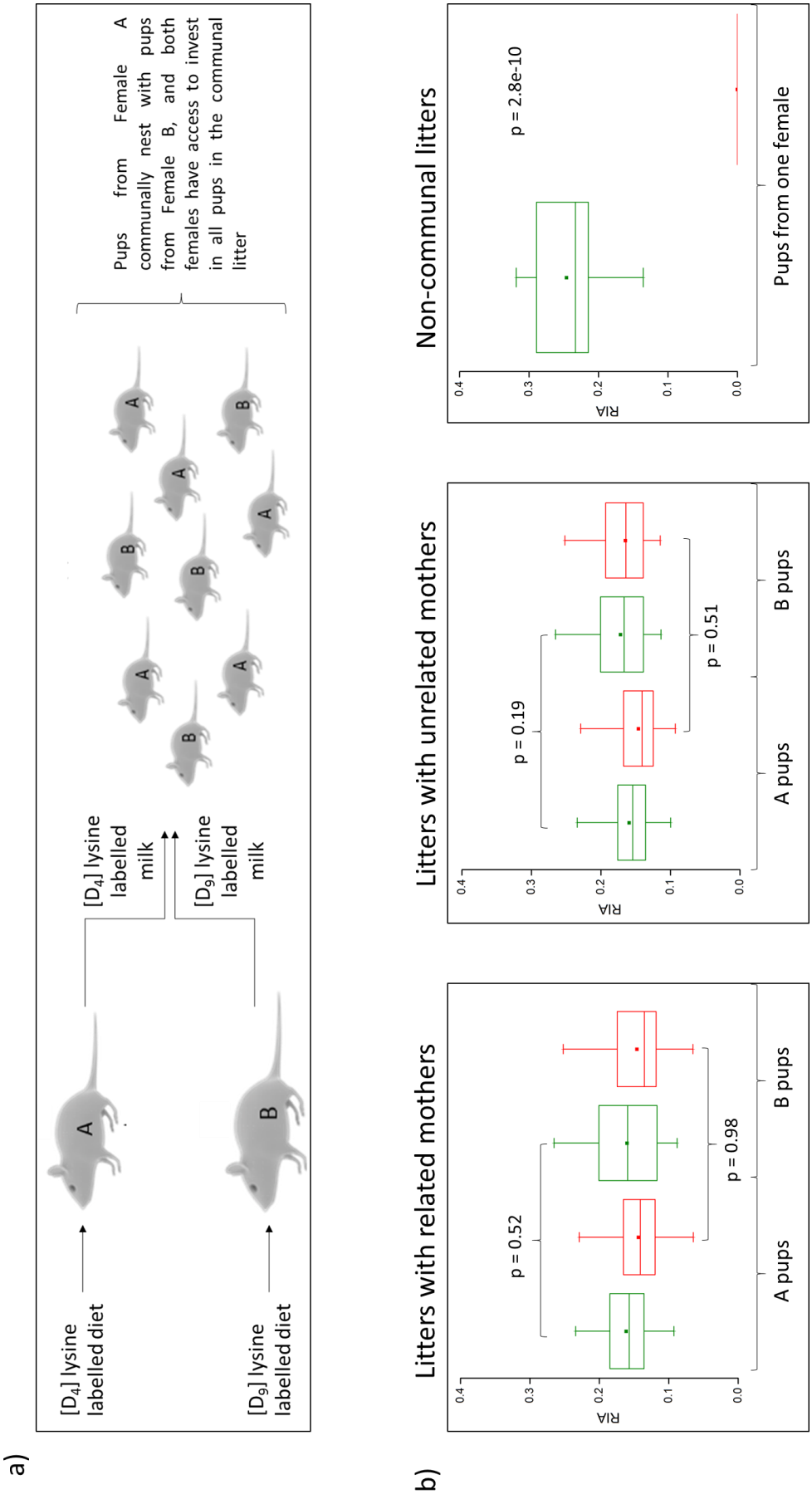
Figure 4.47 Summary of the analysis of pup liver and muscle samples from the milk labelling pilot experiment, using copies of Figure 4.15, Figure 4.16 and Figure 4.21.
(a)The average RIA calculations of the top 20 proteins in pup liver, common in both samples. Pup liver samples taken from two different pups on the final day of the milk labelling experiment were homogenised and digested with Lys-C. The digests were analysed using LC-MS and mass spectra were searched for matches in a mouse database using PLGS. For each of the proteins identified here, average RIAs were calculated from the intensities of the top three peptides identified in each protein, in each sample (error ± SD). (b) The average RIA calculations of the top 20 proteins in pup muscle, common in both samples. Analysis of these samples was the same as for the liver. (c) The rate of heavy amino acid incorporation over 6 days in a high turnover and a low turnover pup liver protein, and a high turnover and a low turnover muscle protein. The plot shows the average daily RIAs (calculated from the top two peptide matches in each protein, from each animal) for a high and low turnover protein in liver (non-specific lipid transfer protein and glutamate dehydrogenase 1, mitochondrial, respectively), and for a high and low turnover proteins in muscle (parvalbumin alpha and beta-enolase, respectively) (error ± SD).

investment over a longer period of time. For each pup tissue type, a protein with a high precursor RIA and one with a low precursor RIA on the final day of the experiment was selected, and the precursor RIAs of these proteins were observed in the tissues over the course of the labelling experiment to determine whether these proteins were high/low turnover and whether they could be used to report variations in labelling at different points during the later communal nursing experiment. Whilst these different proteins incorporated the label at different rates and reached different maximum RIAs by the end of the experiment, indicating that these proteins had different turnover rates, none of the proteins reached a labelling plateau (Figure 4.47). The slow incorporation of label into pup proteins is due their rate of growth, as observed in the analysis of the proteins in pup urine samples. Despite this, all four of these proteins in pup tissues labelled at a significant enough rate to be suitable for analysis of pup tissues to track investment from differently labelled females in the communal nursing experiment.

For the communal nursing study, 8 related female pairs and 9 unrelated female pairs were set up to give birth and nurse their litters in the communal nest. In order to track the investment from each mother in a communal litter, each mother consumed a differently labelled diet (one labelled with [D_4] lysine, the other with [D_9] lysine) via the established feeding mechanism when the first litter in the nest to be born was 7 days old. The mothers were fed these diets for 7 days. This resulted in two different labels being incorporated, one into each mother, which were passed on to the communal litter via their milk, and incorporated into the pups they invested in (Figure 4.48). The precursor RIAs of each label present in the proteins in the pup's tissues indicated the proportion of investment from each mother.

In communal nests where the two mothers were related (7 pairs of females and 7 communal litters), only the females assigned the [D_4] lysine labelled diet appeared fully labelled with the correct diet, with no evidence of consuming any of the unassigned [D_9] lysine diet. Only three out of seven females assigned the [D_9] lysine diet were fully labelled with the correct diet; four were partially [D_9] lysine labelled, and two of these showed evidence of having consumed some of the unassigned [D_4] lysine diet. The reason for these mice having lower [D_9] lysine RIA values is likely to be as a result of them consuming less diet – as urine samples were only taken on the final day of the experiment, it is unknown whether these animals were consuming less diet throughout the experiment, or just on the day prior to sampling (as it had been established in previous experiments that precursor RIAs in MUPs were lower in

Figure 4.48 Summary of the experiment to track investment in communal litters. (a)Pregnant females lived and gave birth in pairs (8 related pairs, 9 unrelated pairs). When both females had given birth, one female was fed [D_4] lysine labelled diet and the other was fed [D_9] lysine labelled diet. These labels were then passed to the pups they invested in via their labelled milk, allowing the investment from each mother to be tracked due to the presence of labels in pup tissue samples. (b) Summary of the RIAs of heavy amino acid incorporation reached after 7 days in litters with related mothers, in litters with unrelated mothers and in non-communal litters, where only the pups from one mother in the pair survived. Mothers were assigned either [D_4] lysine (female A) or [D_9] lysine (female B) labelled diet. Tissue samples were taken from all pups on the final day of the communal nursing experiment, homogenised, digested with Lys-C and analysed using LC-MS. The RIAs in each tissue sample were calculated for glutamate dehydrogenase 1 (low turnover in liver), beta-enolase (low turnover in muscle), parvalbumin alpha (high turnover in muscle) and non-specific lipid transfer protein (high turnover in liver). Boxplots show the calculated RIAs of [D_4] lysine in A pups and B pups (green), and of [D_9] lysine in A and B pups (red) for all related (bottom left) and unrelated (bottom centre) communal litters. The p-values are calculated from Welch t-tests between the [D_4] lysine RIAs calculated for A and B litters in the nests, and the same between the [D_9] lysine RIAs. For the non-communal litters (bottom right), the boxplots show the calculated RIAs of the heavy label received from the mother of the litters (green) and the mother with no surviving pups (red). The p-value is calculated using the Welch two sample t-test.


a)
[D4] lysine labelled diet
[D9] lysine labelled diet
[D4] lysine labelled milk
[D9] lysine labelled milk
A
B
Pups from Female A communally nest with pups from Female B, and both females have access to invest in all pups in the communal litter
b)
Litters with related mothers
Litters with unrelated mothers
Non-communal litters
RIA
p = 0.52
p = 0.98
p = 0.19
p = 0.51
p = 2.8e-10
A pups
B pups
Pups from one female

samples taken after the 'inactive' period). It is unknown whether these mice were consuming less diet because of a possible problem with the feeding mechanism, however this is unlikely due to no problems being encountered in any of the previous experiments. The fact that two of the [D_9] lysine assigned mice had very low intensity [D_4] lysine peptides in their urine could be due to the [D_4] lysine female bringing and dropping small amounts of their [D_4] lysine diet into the communal nest, and so accessible by the [D_9] lysine assigned female. The precursor RIA values for [D_4] and [D_9] lysine labelled peptides were calculated for the four proteins selected in the milk labelling study in the tissue samples of pups from litters where both mothers were fully, correctly labelled. The difference between group mean RIAs and the variation among and between group RIAs was assessed for each of the four proteins using analysis of variance (ANOVA) or in cases where litter sizes were too small for ANOVA, the Welch two sample t-test. To determine whether either female invested significantly more (or less) in the entire communal litter, the difference between entire group mean RIAs was assessed using the Welch two sample t-test. This determined whether there was any statistical significance between the investment from a female in her own pups and in the pups from a related female. There were no statistically significant differences between a mother's investment in her own pups and her sisters in all communal litters (*i.e*. no difference between the [D_4] lysine RIAs calculated in the pups belonging to the [D_4] lysine mother and the pups belonging to the [D_9] lysine mother, and *vice versa*). However, in all litters, one female appeared to invest more in the entire communal litter than the other. The reason for this cannot be determined due to the small number of communal litters studied. There was no apparent link between size or age of litter and the proportion of communal investment from their mother, but when relating back to previous literature, Wilkinson and Baker, (1988) and Konig, (1989) state that in captivity, wild female house mice appear to take turns nursing each other's offspring in communal nests. Taking into account the growth rate and high protein turnover in pups, it is possible that in this communal experiment, the females have been taking turns in indiscriminately investing in the entire communal litter, and it could be that the female who appeared to invest 'more' was in fact the last female to nurse the communal litter before the end of the experiment, and that throughout the experiment, both females were investing similarly, taking it in turns to nurse the entire communal litter.

In communal nests where the two mothers were unrelated (6 pairs of females and 6 communal litters), all females were fully labelled with the correct label, with no evidence of having consumed any of the unassigned diet, which indicates that the

feeding mechanisms are functioning correctly. In four of the six unrelated pairs, no female appeared to discriminate between their own offspring and their non-offspring in terms of investment (*i.e*. no difference between the [D_4] lysine RIAs calculated in the pups belonging to the [D_4] lysine mother and the pups belonging to the [D_9] lysine mother, and *vice versa*). In one pair, there appeared to be a significant difference between the two litters in the communal nest in the [D_4] lysine RIAs calculated in one protein. In another pair, in two proteins, there was a significant difference between the two litters in the [D_9] lysine RIAs calculated. Since these communal litters were large, the differences were not consistent amongst the four proteins, and the differences were not highly significant, it is likely that the differences were due to random sampling and not discrimination in investment by either female. In three of the unrelated female pairs, both females appeared to invest similarly in the entire communal litter. In one pair, one female seemed to invest 'highly' significantly more than the other, and in two pairs, one female invested more in the communal litter. Again, there were no obvious reason for these differences, with no apparent link between size or age of litter and the proportion of communal investment from their mother, but the differences could be due to females taking turns in nursing the entire communal litter.

In the four communal nests where the pups all belonged to only one of the females in the pair, in three pairs, only the mother of the pups in each pair was fully labelled. In one pair, both females were fully labelled, and in all four pairs, no female appeared to consume any of the unassigned diet. The fact that the females with no surviving pups generally appeared partially labelled may be due to them consuming less diet due to stress, or the fact they did not need to consume as much since they were not expending energy investing in the pups. In all four litters, pups received investment from their mothers only, and none from the other female in the communal nest. This was the case in both the related pair and the unrelated pairs. As a result, the differences in investment from each mother in the litter was highly significant. In this study, the proportion of lactation investment from communally nursing house mice in pup litters was successfully studied with the use of two different labels, and it was determined that both related and unrelated female pairs do not discriminate between nursing their own litters and the litter from the other female (Figure 4.48). However, in pairs where only one female's pups survived, only the mother of the litter invested in the pups, with the other lactating female not investing in any of the other female's pups (Figure 4.48). This supports the rejection of the hypotheses that in communal

nests, milk stealing occurs and that pup nursing is indiscriminate (Konig 2006), and also suggests that females alleviating discomfort by ridding themselves of excess milk is not a major reason why female house mice communally nurse. Konig (2006) confirmed that communal nursing increased reproductive success, whether or not females were related, but that mutualistic cooperation was higher in nests where females were related. Konig (1994) also stated that whilst familiar females (regardless of relatedness) nesting communally improved female reproduction and pup survival, female groups which were familiar and related also produced the largest litters with heaviest pups, suggesting that relatedness of females results in increased investment in litters. In this study, both familiar related and familiar unrelated female pairs did not discriminate between their own and other female's pups in terms of nursing, which supports the statements by Konig that increased reproductive success and pup survival results from communal nursing of females, regardless of their relatedness. In the majority of related and unrelated female pairs, one female invested more in the entire communal litter than the other (in all three related pairs and in three of the six unrelated pairs). This indicates no evidence of higher mutualistic cooperation in nests where females are related (Konig 2006). Konig (1994) stated related females produced the largest litters, but this was not found to be the case in this study (Welch two sample t-test: $t = -0.55$, $df = 4.1$, $p = 0.61$). Also, there was no significant difference between the weights of the pups (after the 7 day experiment) in related and unrelated female communal nests (Welch two sample t-test: $t = 0.42$, $df = 101.6$, $p = 0.68$). Manning *et al*. (1995) found that in communal litters where there was an age difference of 3 - 4 days, mothers of the older litters preferentially retrieved their own pups, but the mothers of younger litters did not discriminate between their own offspring and alien offspring. This was not found to be the case in this study, where communal litter age differences ranged from 0 - 5 days, and no female preferentially retrieved their own pups. This was only the case where only one mother had surviving pups. Of interest was the fact that in most communal nests, one mother invested more in the entire communal litter than the other, so it was considered whether litter size or litter age differences linked to the female investing more than the other in the communal litter. In the four nests where the pups born to the [D_9] lysine female were older (age difference ranging from 1-3 days), the [D_9] lysine labelled female invested more in the entire communal litter in three of these nests. In the four nests where the pups born to the [D_4] lysine female were older (age difference ranging 3-5 days), the [D_4] lysine female invested more in the entire communal litter in two of these nests. Therefore, there is no definite link

between age differences and amount of communal investment from a mother. There is also no apparent link between communal investment from a mother and litter sizes within the communal nest, so it is possible that the apparent differences in investment are actually due to the mothers taking turns in investing in the entire communal litter.

Stable isotope labelling strategies and mass spectrometry have enabled investment from communally nursing female mice to be tracked in their pups to discover whether any discrimination in investment is evident in relation to relatedness of the female pairs in the nest. No females, related or unrelated, discriminated between their own pups and their female partners' when investing. As a result, there was no difference in pup weights and litter sizes from related and unrelated female nests; regardless of relatedness, familiar females cooperatively invest in communal litters to increase reproductive success. In order to determine why, in most cases, one female invests significantly more in the entire communal litter than the other, further communal nursing experiments would need to be carried out to discover whether there is a link between investment from a mother and the age of her pups – in more than half of the communal nests, the mother of the older litter invested more, so this is a possible link which is best explored with a greater sample size. To determine whether the differences in investment are observed throughout the communal nursing experiment, or whether the apparent differences in investment are due to mothers taking turns, pup samples would be needed to be removed from the communal nest at intervals during the experiment as well as on the final day.

Chapter 5: Quantification of mouse major urinary proteins

5.1 Introduction

Major urinary proteins (MUPs) in mouse urine play a significant role in olfactory communication amongst conspecifics. MUPs bind, protect and slowly release the volatile components in scent marks that elicit various behavioural and physiological responses in mice, including aggression between male mice (Novotny *et al.* 1985; Jemiolo *et al.* 1985) and the onset of puberty in female mice (Novotny *et al*. 1999). However MUPs, in the absence of volatile pheromonal ligands, also have roles in modulating identity signalling, attractiveness and aggressive responses (Hurst *et al*. 2001; Chamero *et al*. 2007; Roberts *et al*. 2010). Since the regulation of MUP expression of an individual is altered by situations such as reproductive status and social setting (Stopka *et al.* 2007), and that in wild mice, MUP profiles could be used by conspecifics in recognition and to avoid inbreeding, it is essential that as well as identifying the MUPs and their roles in chemical signalling, the differences in MUP expression profiles are quantified. For a definitive assessment of the variation in MUP expression of an individual in different situations, and the differences in MUP expression profiles between individuals, a suitable quantification method needs to be developed.

Previously, the absolute quantification of MUPs was based on QconCAT technology; a label–mediated methodology for absolute protein quantification. A QconCAT is an artificial concatamer of different peptides from a number of different protein targets. This [$^{13}C_6$] lysine labelled standard is added to the sample prior to proteolysis, and upon digestion, the labelled 'Q peptides' are generated, which act as standards for the target peptides in the sample in LC-MS analysis (Beynon *et al*. 2005; Pratt *et al*. 2006; Brun *et al.* 2009). Quantification of analyte peptides is by the ratio between the analyte ('light' peptides) and previously-quantified QconCAT ('heavy') peptides (Pratt *et al*. 2006). A QconCAT for the quantification of the inbred laboratory strain C57BL/6 MUPs was designed by Dr S. Armstrong and Dr. D. Simpson, University of Liverpool, and the quantification method was developed and implemented by Dr. J. Unsworth (University of Liverpool). However, problems were encountered using a QconCAT for the absolute quantification of MUPs (Beynon *et al*. 2014; 2015). Firstly, the high sequence similarity between the MUP variants, particularly central MUPs, meant that for some MUPs, no unique peptide could be generated for use in the quantification of

that protein. Other problems encountered included peptides undergoing deamidation, incomplete digestion of MUPs and peptides that ionised poorly in mass spectrometric analysis (Dr. J. Unsworth, Ph.D Thesis). Despite these challenges, the QconCAT methodology was successfully implemented in the quantification of MUPs from the inbred C67BL/6 laboratory strain (Beynon *et al*. 2015).

Whilst the QconCAT was effective for the quantification of the MUPs expressed by an inbred laboratory strain, further complications could arise from the high rate of evolution of these proteins (Beynon *et al.* 2014); wild–caught mice have been found to express new allelic variants of MUP (Robertson *et al*. 1996; Beynon *et al*. 2002) which upon proteolysis may generate peptides different to those in the designed MUP QconCAT. The analysis of MUPs from different *Mus* subspecies may also reveal MUPs that are different in sequence to those present in the C57BL/6 laboratory mouse strain (Robertson *et al*. 2007), and in both of these cases, new QconCATs would have to be designed and constructed. Given the challenges in the design of the existing MUP QconCAT, the development of a new method that is more suitable for the absolute quantification of highly homologous proteins to quantify new MUP variants is essential, and this method could be used alongside the existing QconCAT methodology that quantifies the known C57BL/6 MUPs.

A popular method for the absolute quantification of proteins involves the use of isotope labelled full length proteins known as PSAQ standards, with the term ‘PSAQ’ being an abbreviation of ‘Protein Standards for Absolute Quantification’ (Brun *et al.* 2007) (Figure 5.1). As with QconCATs, PSAQ standards share the same biochemical properties as their targets, and so behave identically to the proteins in the sample throughout sample preparation and MS analysis (Dupuis *et al*. 2008; Brun *et al*. 2009). Also, because these standards are spiked in defined amounts at the beginning of sample preparation, there are no differences between standard and sample digestion efficiency, providing more accurate quantification at MS level (Kaiser *et al*. 2011). PSAQ provides greater sequence coverage due to the use of a full length protein standard, meaning isoforms and variants may also be distinguishable (Brun *et al*. 2009). Since the current challenges with MUP quantification are mainly based on the proteins’ sequence homology, the high rate of MUP evolution and the lack of a unique peptide for each protein, the use of a full length protein standard, based on the idea of the PSAQ methodology, may address these problems. Instead of using the ‘bottom up’ workflow employed with PSAQ standards prior to MS analysis (Brun *et al*. 2007), it was considered whether quantification could take place at the intact

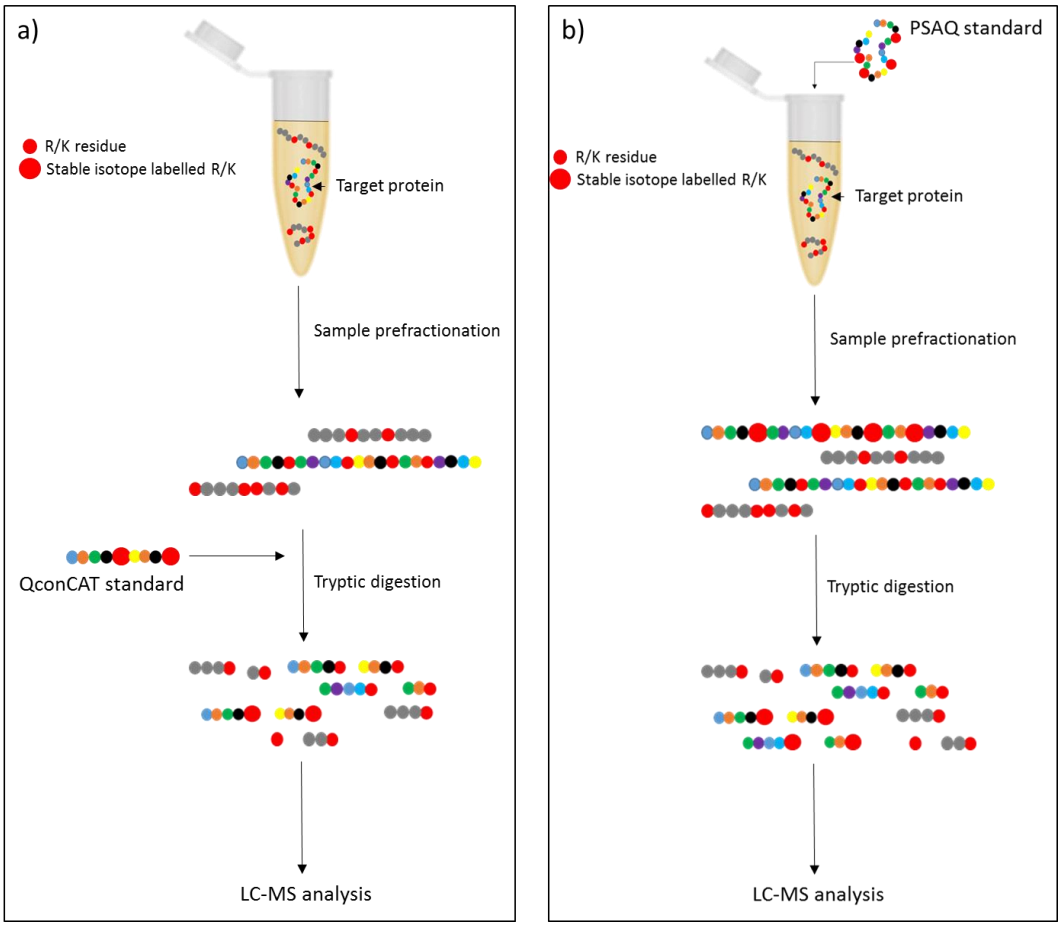


Figure 5.1 Summary of a) QconCAT and b) PSAQ quantification workflows.
a) MS–based absolute quantification of proteins using a QconCAT, an artificial concatamer of different isotope-labelled proteotypic peptide(s) of the target protein(s). This is added prior to digestion, where the labelled 'Q' peptides are generated. b) MS–based absolute quantification using PSAQ, where an isotope labelled version of the target protein is added directly to the sample at the beginning of sample preparation.

protein level ('top down'), since intact protein analysis, by electrospray ionisation mass spectrometry (ESI-MS), is already well established for the accurate identification of MUP isoforms (Evershed *et al*. 1993; Robertson *et al*. 1996; Beynon *et al*. 2002; Robertson *et al*. 2007).

For the analysis of MUPs, ESI-MS has proved invaluable. The urine of an individual mouse contains a number of MUP variants, and MUP expression can vary greatly in individuals and amongst wild-caught mice; ESI-MS is able to accurately identify and distinguish between the molecular weights of most MUP isoforms, which has allowed the comparison of MUP phenotype between mice of inbred strains and wild populations (Robertson *et al*. 1996; Beynon *et al*. 2002; Robertson *et al*. 2007). Whilst the processed and deconvoluted spectra from ESI-MS analysis of MUPs primarily reports the molecular weights of the MUPs present in the sample, it can also provide an indication of the amount of each MUP present in the sample, relative to each other. MaxEnt is a method which 'simplifies the electrospray spectra from complex mixtures' based on maximum entropy, which measures molecular weight and quantitative relative intensities of the sample components' (Ferrige *et al*. 1991; Ferridge *et al*. 1992; Cottrell and Green 1993; Cottrell and Green 1998). Maximum entropy techniques, having previously been applied to Raman and NMR spectroscopy for signal to noise enhancement and spectral deconvolution, were first successfully applied to electrospray mass spectra in 1991 (Ferridge *et al.* 1991). Electrospray spectra of protein mixtures are complex, with each protein in the mixture represented by a multiply charged ion series (Ferridge *et al.* 1992). Therefore, to aid the interpretation of mass spectra of protein mixtures, a way to transform each component from the multiply charged ion series to a single peak on a molecular mass scale is required. MaxEnt is a deconvolution method based on probability, that 'repeatedly processes different trial mass and charge spectra, comparing the results with the observed data' (Ferridge *et al.* 1992). The molecular weights of the sample components are generated from the MaxEnt result; the most probable result (MaxEnt spectrum) is presented from the probability distribution (MaxEnt result), which is calculated from trial spectra that both agree well with the observed data and have large entropy (Ferridge *et al.* 1992). The quantitative data is generated from the summation of the peak intensities from the original multiply charged spectra, and this is represented by the area under the molecular weight peak in the processed MaxEnt spectra (Cottrell and Green 1998).

The MaxEnt software alone cannot be used to give an accurate representation of the amount of each MUP isoform present in a sample. This is because although MUPs are very similar in sequence, they may ionise differently to each other in the source during ESI-MS due to differences in charge state distribution, charge accessibility and hydrophobicity (Grandori 2003; Heck and Van Der Heuvel 2004), and differences in ionisation efficiency may affect the accuracy in the resulting relative quantitation data. For this reason, adding a standard that has identical biochemical properties to its target but with the addition of a label in order to distinguish it from the analyte in mass spectrometry analysis, in a known amount to a sample prior to intact mass analysis by ESI-MS would be desirable, in conjunction with a standard curve. For the quantification of MUPs, recombinant forms of each protein would be ideal as full-length protein standards, as these are the same structurally and in amino acid sequence (except for the addition of the hexa-histidine tag on the N terminal) as their target proteins whilst remaining distinguishable by ESI-MS due to presence of the N terminal hexa-histidine tag, which is added to the recombinant protein for purification.

The aim of this chapter is to conduct a preliminary investigation as to whether ESI-MS analysis of MUPs, in conjunction with the use of known amounts of recombinant MUPs (rMUPs) as standards, can allow absolute quantification of MUPs, using the measurement of the area under curve of the internal standard and the corresponding analyte protein. ESI-MS analysis (using Micromass Q-ToF micro and Waters Synapt G1 Q-ToF mass spectrometers) of mouse urine has confirmed that the peak areas of MUPs (in the true mass spectra) increase linearly with increasing MUP load, and therefore intact mass analysis has been used to provide relative quantification of MUPs in mouse urine using MaxEnt (Mudge *et al.* 2008; Beynon *et al*. 2015; Sheehan *et al.* 2015). However, due to the use of a new Q-ToF mass spectrometer in the laboratory for ESI-MS analysis of intact proteins, it was important that ESI-MS linearity of MUPs was revisited to confirm the data generated by the new instrument was quantitative. The ionisation efficiencies of MUPs relative to one another in samples of differing complexity were also assessed to confirm that ESI-MS analysis of intact proteins was also suitable for the relative quantification of MUPs in urine samples. To do this, the predicted ionisation efficiencies of MUPs, based on differences in their number of protonatable sites, hydrophobicity and charge state distribution profiles, were assessed. Recombinant versions of three of the MUP isoforms present in the male C57BL/6 urine (MUPs 7, 11 and 20) were subjected to ESI-MS analysis as single-protein samples to determine the ionisation efficiencies of MUPs relative to one another. This confirmed whether differences in the ionisation efficiencies exhibited

by each MUP relative to each other could be accounted for by differences in the number of protonatable sites present and/or their charge state distributions, enabling ESI-MS responses for each MUP to be definitively determined, essential if rMUPs were to be used as standards for the absolute quantification of MUPs in urine. ESI-MS analysis of recombinant MUPs as single-protein samples and as part of equimolar mixtures of different complexities (*i.e* two-protein and three-protein equimolar mixtures), plotting the peak areas of each rMUP at different total protein loads, determined whether the ESI-MS responses and ionisation efficiencies of rMUPs as part of more complex samples remained the same, confirming whether ESI-MS analysis of intact MUPs was a suitable method for absolute quantification. ESI-MS responses of rMUPs (and MUPs) would be required to be the same in less complex (*e.g*. female C57BL/6 urine) samples as in more complex samples (*e.g*. the urine of wild-caught male house mice), to allow absolute quantification to take place at the intact protein level without the need for extensive chromatographic separation of urine prior to ESI-MS analysis.

5.2 Results

5.2.1 Predicted ionisation efficiencies of MUPs

ESI-MS has been instrumental in the identification and calculation of relative abundance of MUPs in mouse urine, allowing the accurate and precise determination of the molecular weights of these proteins. ESI of a mouse urine sample, containing MUPs, generates a series of multiply charged protein profiles, which is subsequently deconvoluted using MaxEnt to generate a true mass spectrum from which the average mass of each of these proteins are calculated. The urine of the inbred C57BL/6 mouse strain is commonly used for the study of MUPs, since sequencing of this genome is complete and the MUP locus has been subjected to in depth gene analysis (Mudge *et al*. 2008; Logan *et al*. 2008).

Despite ESI-MS being a powerful tool for the identification of MUPs in a sample through the determination of their molecular weights, limitations arise due to the high homology of MUPs. For example, MUPs with masses of 18693 Da and 18694 Da are often unable to be resolved and identified by mass spectrometry – highlighted in red in Table 5.1 are the MUPs with these masses that are therefore usually indistinguishable. MUP masses 18708 Da and 18713 Da may also be incompletely resolved, so that although these two different masses may be distinguished, calculation of their relative abundances may be inaccurate.

Other limitations of ESI-MS are related to the calculations of relative abundances of MUPs in a sample. Whilst the area under the molecular weight peak in the true MaxEnt ESI-MS mass spectrum generally provides a relative concentration of that protein in the sample, it must be remembered that the differences between the ionisation efficiencies (the ratio of the number of ions formed to the number of molecules in the ion source) of different proteins in a sample means that the amount of each protein detected by the mass spectrometer may not be truly representative of the amount of protein in the sample. For example, a protein that ionises less efficiently is less likely to successfully ionise and therefore be subjected to mass spectrometry analysis than a protein that ionises well. Since MUPs are very similar in amino acid sequence and structure, it is expected that these proteins will have similar ionisation efficiencies, but differences in the number of protonatable sites, hydrophobicity and charge state distribution must be considered to determine whether ESI-MS analysis can be successfully used for MUP quantification.

Protonatable sites

The protonatable sites in a protein are the amino acids with positively charged side chains, and these are histidine, lysine and arginine. The sequences of each MUP were studied for the presence of these amino acids, and the instances of these amino acids were identified (Figure 5.2) and summed (Table 5.1). Whilst the number of protonatable sites in the highly homologous central MUPs were fairly similar (ranging from 23 to 25 protonatable sites), the number of protonatable sites in the less similar peripheral MUPs ranged from 18 to 28. Evidence indicates that the number of protonatable sites a protein has is a major factor in determining the charge state distribution profiles for denatured proteins in ESI (Krusemark *et al.* 2009), so proteins with more protonatable sites are more likely to exhibit a higher charge state distribution, and thus ionise more efficiently in the ESI source.

Hydrophobicity

The number of amino acids with hydrophobic side chains in a protein will affect overall hydrophobicity of the protein – these amino acids are alanine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine and tryptophan. Increased hydrophobicity increases ion yield in positive-ion ESI and therefore increase signal intensity (Fenn 1993; Cech *et al.* 2001; Null *et al*. 2003; Osaka and Takayama 2014) In electrospray ionisation, analytes with a greater hydrophobicity will partition to the surface of the hydrophilic electrospray droplet, thus tending to carry a greater proportion of the charge produced in the ionisation process (Cech and Enke 2001). The sequences of each MUP were studied for the presence of these hydrophobic amino acids, and the highly homologous central MUPs all have a total of 58 hydrophobic amino acids in their sequences. The sequences of each peripheral MUP contained slightly more hydrophobic amino acids, ranging from 62 in MUP 5 to 67 in MUP 21. The similarity in hydrophobicity amongst all MUPs suggests that, in the ESI-MS analysis of MUPs, hydrophobicity is unlikely to be a factor in any differing ionisation efficiencies observed.

Charge state distribution

ESI of denatured proteins generates a series of multiply charged ions which reflect the numbers of protons binding to different sites on the protein (Beynon *et al.* 2014). Proteins with a higher charge state distribution ionise more efficiently in the ESI source than those with a lower charge state distribution, and those with a lower charge state distribution may generate a lower-intensity mass spectrometry signal

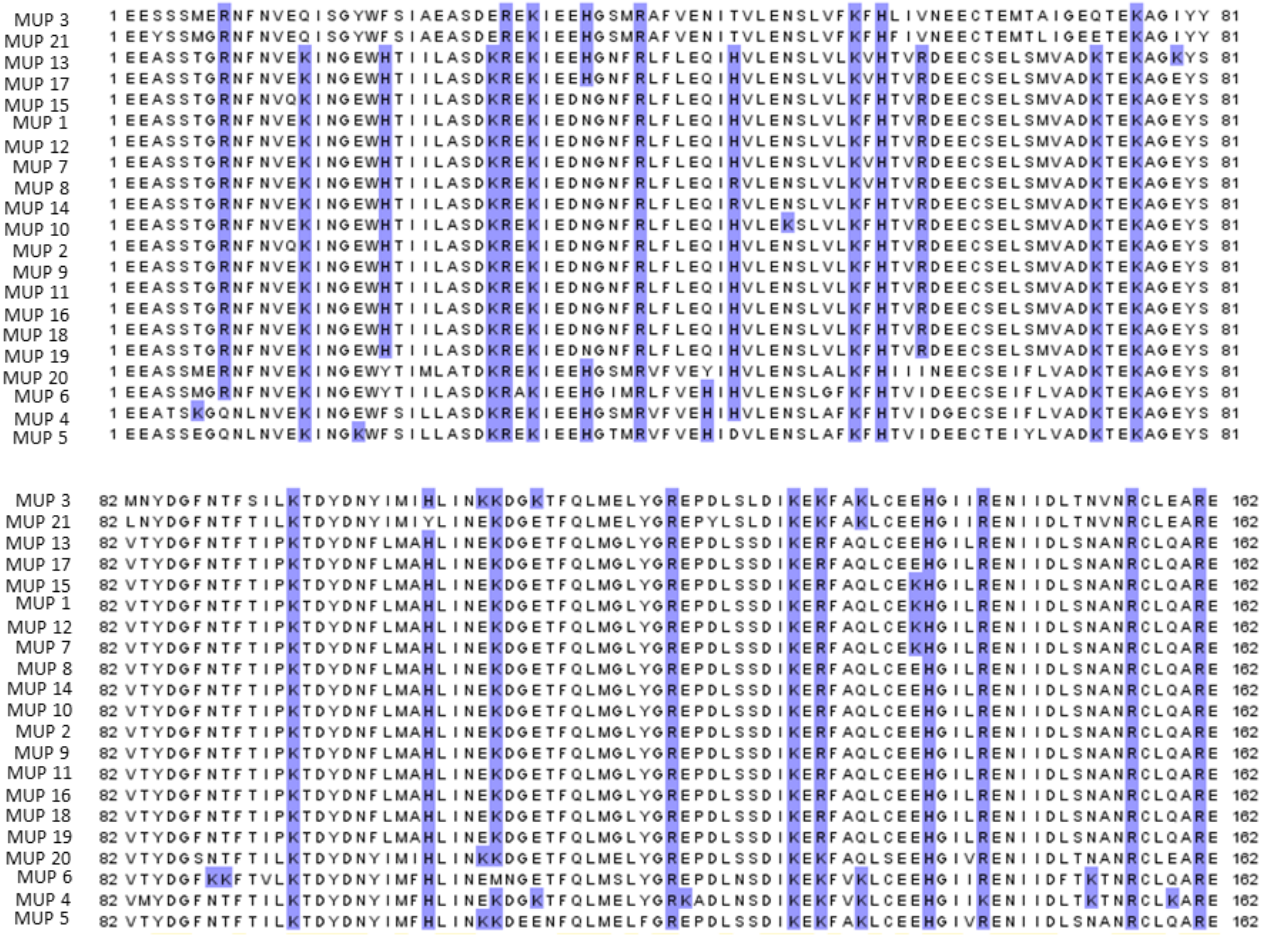


Figure 5.2 Identification of protonatable sites in MUP sequences.
All mature MUP sequences were aligned using Clustal Omega (www.ebi.ac.uk) and the multiple alignment was viewed using Jalview. Histidine, lysine and arginine residues (protonatable sites) are highlighted in blue.

Table 5.1 MUP summary.
Included in the summary are the MUP names, their mature masses, whether they originate from the central or peripheral region of the MUP gene, and the number of protonatable sites each MUP sequence possesses. Highlighted in orange are the MUPs with masses of 18693 Da and 18694 Da, and so are indistinguishable in intact protein analysis.

MUP name	Mature MUP mass (Da)	MUP gene region	Number of protonatable sites
MUP 1	18693	Central	24
MUP 2	18693	Central	23
MUP 3	18956	Peripheral	21
MUP 4	18816	Peripheral	28
MUP 5	18863	Peripheral	18
MUP 6	18985	Peripheral	26
MUP 7	18645	Central	24
MUP 8	18665	Central	23
MUP 9	18694	Central	23
MUP 10	18708	Central	24
MUP 11	18694	Central	23
MUP 12	18693	Central	24
MUP 13	18682	Central	25
MUP 14	18713	Central	23
MUP 15	18692	Central	24
MUP 16	18694	Central	23
MUP 17	18683	Central	24
MUP 18	18694	Central	23
MUP 19	18694	Central	23
MUP 20 (Darcin)	18893	Peripheral	23
MUP 21	19109	Peripheral	18

than expected. In the ESI-MS analysis of MUPs, the charge state distribution profiles are usually distributed around the $[M+17H]^{17+}$ ion. Despite the high similarity of MUPs, which would indicate that their charge state distributions would be very similar, MUP 20 (darcin) exhibits a different, lower charge state distribution than other MUPs (Phelan *et al.* 2014). This is consistent with MUP 20 retaining a more compact structure that is not completely unfolded under the conditions used for ESI-MS analysis, further confirmed by the fact MUP 20 migrates at a higher mobility than other MUPs on SDS-PAGE (Phelan *et al.* 2014).

The similarity of MUP sequences means that there are unlikely to be any major differences in ionisation efficiencies amongst them due to differences in hydrophobicity. The similarity of MUP sequences and structure indicate that there are also unlikely to be any notable differences in ionisation efficiencies due to differing charge state distributions in ESI-MS analysis. The exception is MUP 20, which exhibits a lower charge state distribution profile, suggesting this MUP may generate a lower intensity mass spectrometry signal, and so differences in charge state distribution profiles will be considered in future analyses. The highly homologous central MUPs also have a very similar number of protonatable sites, again suggesting no differences in ionisation efficiencies will be observed. However, the peripheral MUPs possess more variable numbers of protonatable sites, which may result in differences in charge state distributions and therefore ionisation efficiencies. Since evidence indicates that the number of protonatable sites a protein has, rather than its shape and molecular mass, appears to determine ionisation efficiency for denatured proteins (Krusemark *et al.* 2009), this will be another consideration when assessing the relationship between MUP concentration and observed ESI-MS response, and whether ESI-MS analysis can be used as a tool for accurate quantification of MUPs.

5.2.2 ESI-MS linearity of MUPs and rMUPs using the Waters Synapt G1

The true MaxEnt ESI-MS mass spectrum of male C57BL/6 urine contains four predominant masses, which are 18645 Da, 18694 Da, 18708 Da and 18893 Da. All these masses match those of previously identified MUPs, with the 18645 Da mass representing MUP 7, the 18708 Da mass representing MUP 10 and 18893 Da representing MUP 20 (darcin). The 18694 Da mass represents one of the following: MUPs 1, 2, 9, 11, 12, 16, 18, 19 (Table 5.1; Figure 5.3). The MUPs 9, 11, 16, 18 and 19 have identical amino acid sequences and therefore identical mature masses of 18694 Da. MUPs 1, 2 and 12 have an identical mature mass of 18693 Da, and whilst MUPs 1 and 12 have identical amino acid sequences, MUP 2 has this mass from a

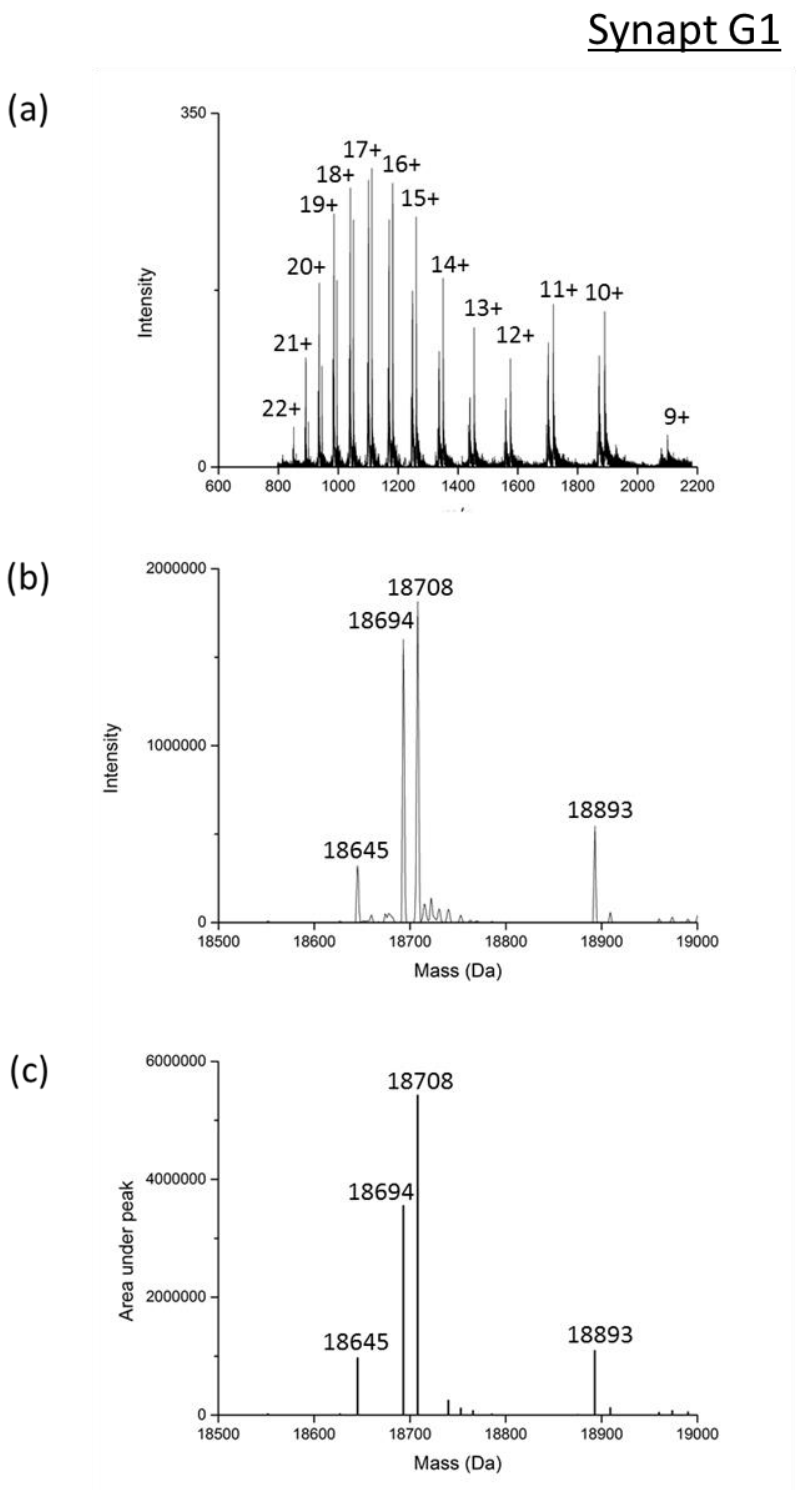


Figure 5.3 ESI-MS analysis of male C57BL/6 urinary MUPs.
Urine (diluted in 0.1% (v/v) formic acid) was injected onto a C4 desalting trap and the masses of the MUPs present were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). a) The multiply charged ion series for each MUP. b) The deconvoluted, true mass spectrum, showing the intact mass for each MUP. c) The true mass spectrum viewed as peak areas, which give the relative concentrations of each MUP in the sample.

different primary sequence. The 18693 Da mass cannot be distinguished from 18694 Da in the ESI-MS analysis of intact proteins, and so a mass of 18694 Da in an ESI-MS spectrum could arise from any of these MUPs. In female urine, the true ESI-MS mass spectrum contains two of these dominant masses, 18694 Da and 18708 Da (Mudge *et al.* 2008) (not shown). Previous work by Cheetham *et al.* 2008 and Mudge *et al.* 2008 confirms that these MUP expression patterns are consistent amongst individuals of the same sex of this strain.

If ESI-MS is to be used as a quantitative method for the analysis of MUPs, there needs to be a linear relationship between MUP in a urine sample and the area under MUP peak in the deconvoluted, processed true mass spectrum. Previous intact MUP ESI-MS analysis undertaken by Dr. S. Armstrong, University of Liverpool, (Ph.D Thesis) explored the relationship between protein load and ESI-MS response, however this experiment was based on the peak intensities of the MUPs in the true mass spectra rather than the peak areas. The peak intensity is less reliable as an indication of relative MUP intensity; it is conditioned by the software's confidence that the reported molecular weight is the correct molecular weight. It does give an insight into the possible relative amounts of MUP present, since the software will be more confident of assigning the correct molecular weight to a more abundant protein in a sample than it will be to a less abundant protein. As described earlier, it is the area under the MUP peaks, generated from the summed intensities of all the multiply charged ions present in the mass spectrum, that gives relative quantification of the MUP in that sample.

To assess the relationship between MUP concentration and MUP peak area, C57BL/6 male urine was diluted to a total protein concentration of 1.6 ng/µl. Diluted urine (0.5, 1, 1.5, and 2 µl) was injected into the system and ESI-MS analysis was carried out using the method described in Chapter 2, using a Waters Synapt G1 Q-ToF mass spectrometer. The raw mass spectra were processed as described in Chapter 2, giving the molecular weight of each MUP and their peak intensities (Figure 5.3). The area under peaks (and therefore relative MUP concentration) increased with increasing amounts of MUP on column from 0.8 ng on column up to amounts of 3.2 ng (Figure 5.4). Each sample was run in triplicate, with minimal variation observed in the values obtained for the under-peak areas (Figures 5.4 and 5.5). For each MUP, the linear relationship was tight, with r^2 values of 0.99 for all MUPs in the sample. Whilst the relationships between protein load and area under peak of MUP 11 (18694 Da) and MUP 20 (18893 Da) are very similar (suggesting a similar relative amount of

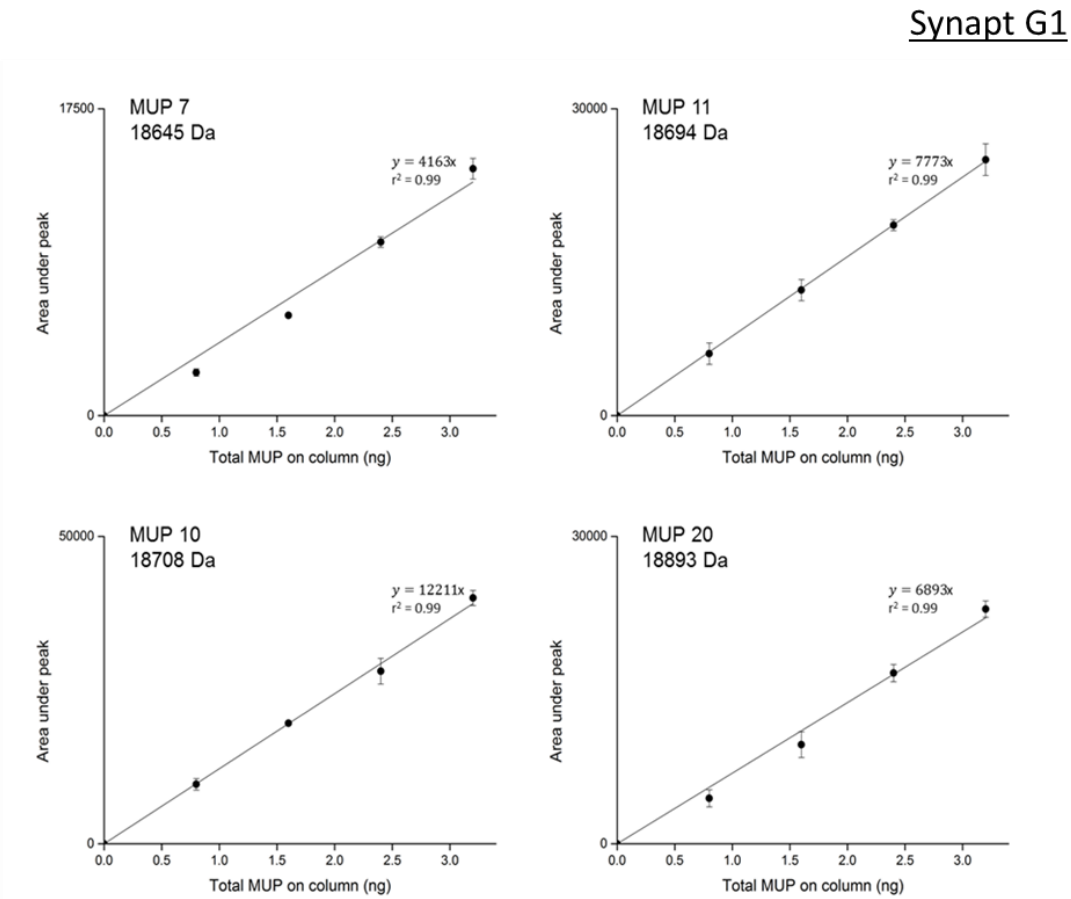


Figure 5.4 Linearity of ESI-MS peak area with increasing protein load for each MUP in the male C57BL/6 urine sample.
Urine (diluted in 0.1% (v/v) formic acid) was injected onto a C4 desalting trap and the masses of the MUPs present were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). Samples were injected in triplicate. The area under peak values for the four main male C57BL/6 MUP variants were assessed as total MUP load on column increased. Error bars represent SD (n = 3).

Synapt G1

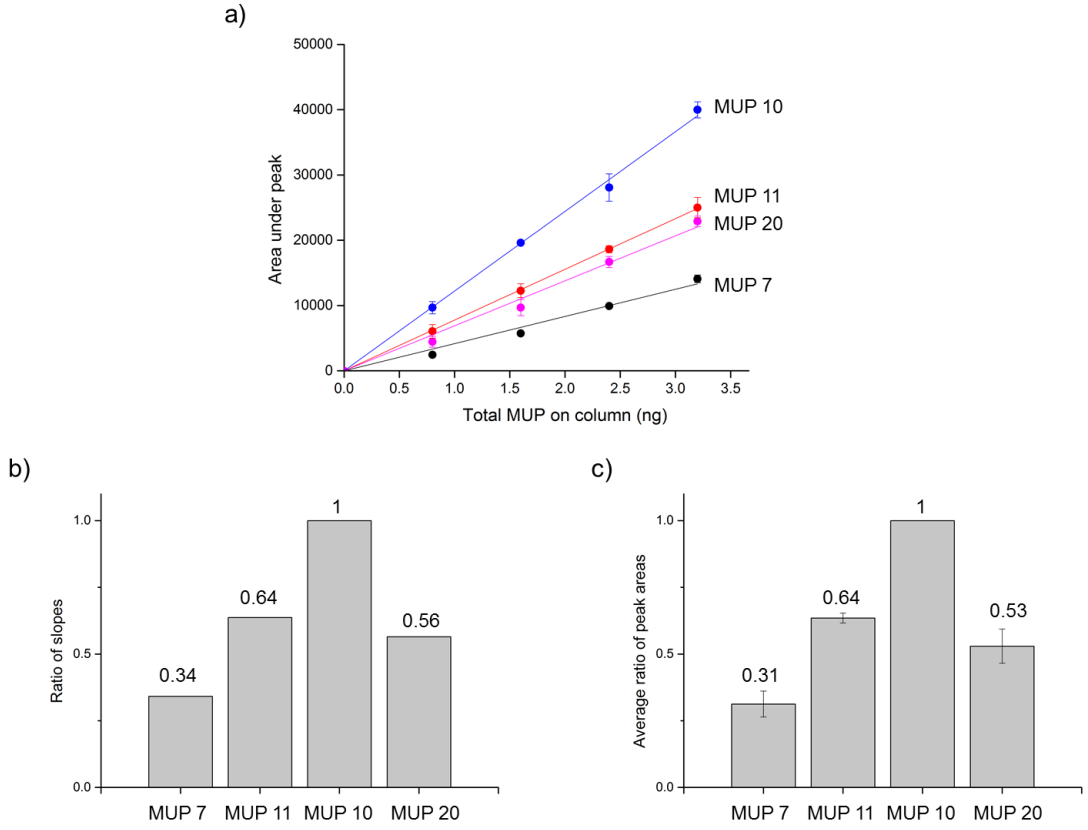


Figure 5.5 Comparison of the linearity of MUP ESI-MS peak area with increasing protein load.
a) The area under peak values for the four main male C57BL/6 MUP variants assessed were plotted on the same scale for a comparison between the peak area/protein load relationships exhibited by each of the MUPs in the urine sample. Error bars represent SD (n = 3). b) The ratio of the slopes of the linear relationship for each MUP, normalised to 1 (the value of the steepest gradient). c) The average ratio of peak areas for each MUP normalised to 1 (the largest MUP peak area at each protein load). n = 4 (each protein load), error ± SD.

these proteins is present in the sample), the linear relationship for MUP 7 (18645 Da) is shallower (indicating there is less of this protein in the sample relative to the other MUPs), and the linear relationship for MUP 10 (18708 Da) is steeper (suggesting there is more of this MUP in the urine sample relative to all other MUPs in the urine sample). The linear ESI-MS responses of all MUPs in the sample allowed the relative quantification of these proteins, as displayed in the ratios of linear slope values and the ratios of MUP peak areas at different protein loads in Figure 5.5, under the assumption that all MUPs in the sample have similar ionisation efficiencies.

Upon assessment of the multiply charged ion series generated by the mixture of proteins, it was noted that two different charge state distribution profiles were present; a major distribution centred around the $[M+17H]^{17+}$ ion (the usual distribution observed in the analysis of MUPs), and a second, less intense distribution centred around the $[M+11H]^{11+}$ ion (Figure 5.6). This was not observed in the analysis of horse heart myoglobin (analysed for instrument calibration), which exhibited one charge state distribution around the $[M+21H]^{21+}$ ion, indicative of the protein being fully denatured under these experimental conditions (Figure 5.6). The higher charge state distribution observed for horse heart myoglobin compared to the MUPs in male C57BL/6 urine is due to it having a greater number of protonatable sites (horse heart myoglobin has 32 sites, MUPs 7 and 10 have 24 and MUPs 11 and 20 have 23). The fact that two different charge state distribution profiles are present in the ESI-MS analysis of mouse urine suggests that at least one of the MUPs in the sample is present in two different conformational states, with the less intense charge state distribution profile resulting from a protein that is retaining some folding in the gas phase, which has fewer basic sites available for protonation. Closer analysis of the multiply charged ion series indicated that all proteins in the sample exhibited a second distribution centred around the $[M+11H]^{11+}$ ion, meaning that all four major MUP isoforms retained some structure in ESI-MS analysis.

It was required that recombinant MUPs, the intended protein standards for absolute quantification of MUPs, also exhibited a linear relationship between protein load and peak area. Since these proteins are identical in sequence to their native counterparts (except for the addition of the hexa-histidine tag used for purification), it would be assumed that this would be the case, however, due to native MUPs in male C57BL/6 urine retaining some folding in the gas phase, it was essential to confirm that the recombinant forms of these proteins exhibited similar conformational states under the same experimental conditions. Recombinant forms of MUPs 7, 11 and 20 (darcin)

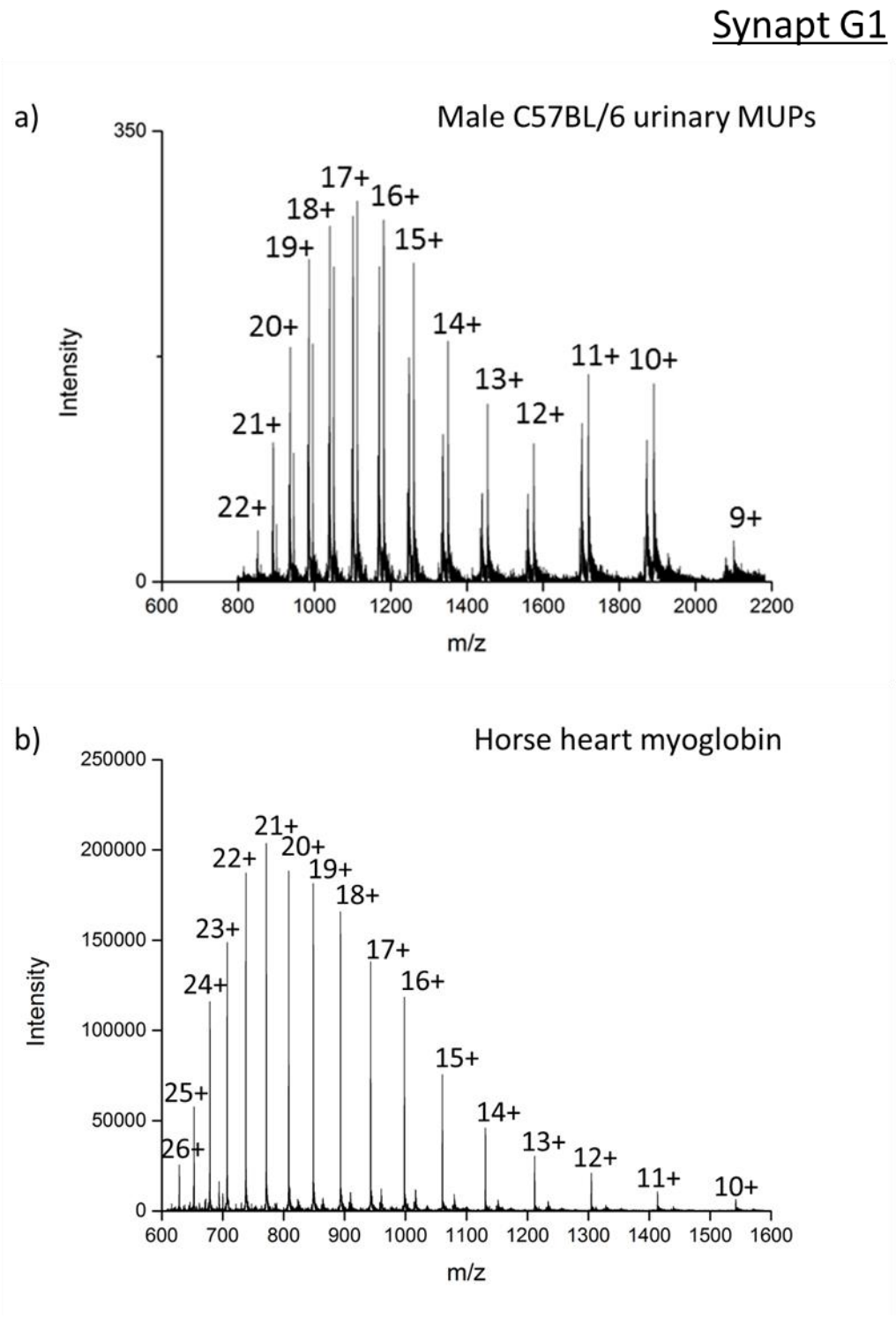


Figure 5.6 The charge state distribution profiles for a) a MUP mixture and b) horse heart myoglobin.
a) The multiply charged ion series for MUPs in a male C57BL/6 urine sample, with the two distinct distribution profiles indicating the proteins are retaining some structure under these experimental conditions. b) The multiply charged ion series for horse heart myoglobin, used as an external calibrant for the experiment. This protein, under the same conditions, exhibits one, positively skewed normal distribution profile, meaning this protein is fully denatured in the ESI source.

with an N terminal hexa-histidine tag were expressed by Mrs L McLean as described in Chapter 2, and were purified using affinity chromatography, also described in Chapter 2. The concentration of each protein was confirmed by protein assay - MUP 7 (18645 Da) was 4.3 mg/ml; MUP 11 (18694 Da) was 5.4 mg/ml and MUP 20 (18893 Da) was 14.4 mg/ml. The rMUP samples were diluted to approximately 2 ng/µl, and 0.5, 1, 1.5, and 2 µl of each rMUP sample was injected into the system and ESI-MS analysis was carried out using the method described in Chapter 2, using a Waters Synapt G1 Q-ToF mass spectrometer. All injections were analysed in duplicate.

Whilst the charge state distribution profiles in ESI-MS analysis are usually centred on the $[M+17H]^{17+}$ ion for natural MUPs, the charge state distribution profiles in the ESI-MS analysis of rMUPs are different; centred around the $[M+20H]^{19+}$ ion for rMUP 20 (darcin) and around the $[M+21H]^{21+}$ ion for rMUPs 7 and 11 (Figure 5.7). The higher charge states observed in the ESI-MS analysis of recombinant MUPs compared to natural MUPs is due to the addition of further protonatable sites to each protein in the form of the N-terminal hexa-histidine tag. The lower charge state distribution observed in the ESI-MS analysis of rMUP 20 was observed as expected from previous analysis of natural MUP 20 (Phelan *et al.* 2014), confirming that both natural and recombinant forms of the protein are retaining a more compact structure in ESI-MS analysis. All three rMUPs exhibited two conformational states, with a second distribution profile centred around the $[M+11H]^{11+}$ ion, suggesting that they all retain some degree of folding under these experimental conditions, as seen in the analysis of MUPs in male C57BL/6 urine.

Raw mass spectra were processed as described in Chapter 2, giving the molecular weight of each rMUP and their peak intensities (Figure 5.7). The area under peaks (and therefore relative rMUP concentration) increased linearly with increasing amounts of protein on column (Figure 5.8), and for each rMUP, the linear relationship was tight, with r^2 values of 0.99 for all proteins. For ESI-MS analysis of MUPs to be a suitable method for absolute quantification, knowledge of the responses of each individual MUP in ESI-MS analysis is needed, involving assessment of the relationship of MUP concentration and MUP peak area whilst considering how charge state distribution, protonatable sites and sample complexity may affect MUP ionisation and peak area. This was explored using the recombinant forms of MUPs, which was deemed suitable for the analysis of individual MUPs without the need for complex and time-consuming chromatographic separation of mouse urine.

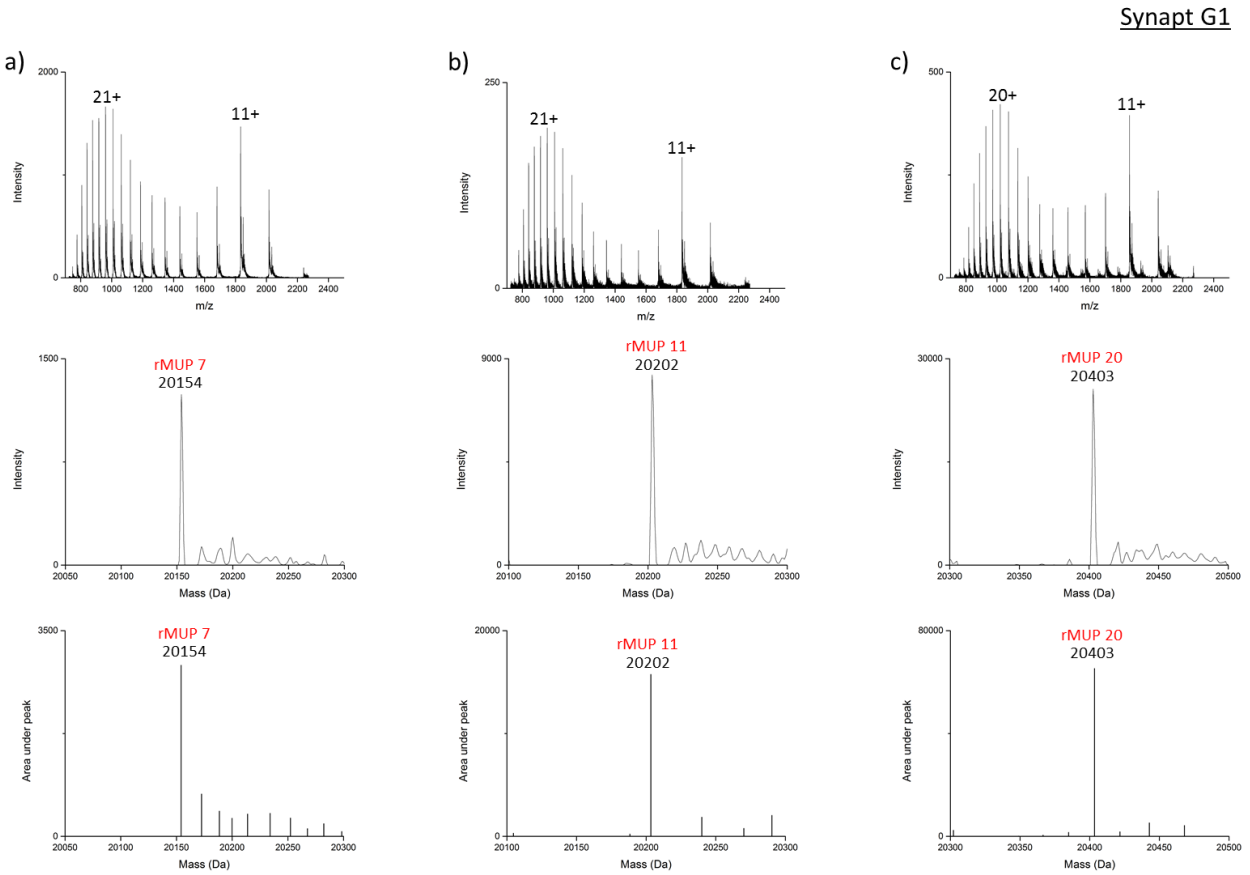


Figure 5.7 ESI-MS analysis of separate rMUP samples using the Synapt G1.
Samples (diluted in 0.1% (v/v) formic acid) were injected onto a C4 desalting trap and the masses of the rMUPs were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). a) The multiply charged ion series, true mass spectrum and area under peak for rMUP 7, b) The multiply charged ion series, true mass spectrum and area under peak for rMUP 11, and c) The multiply charged ion series, true mass spectrum and area under peak for rMUP 20.

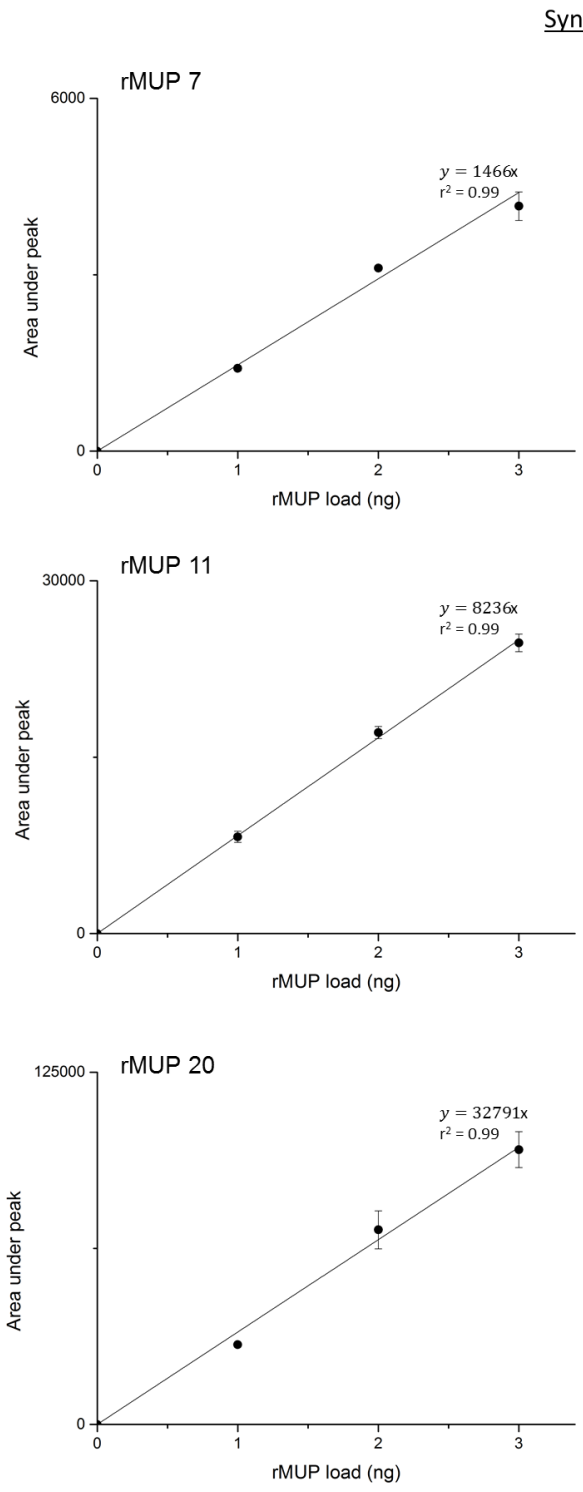


Figure 5.8 Linearity of ESI-MS peak area with increasing protein load for each rMUP using the Synapt G1.
Samples (diluted in 0.1% (v/v) formic acid) were injected onto a C4 desalting trap and the masses of the MUPs present were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). Samples were injected in duplicate. The area under peak values for the rMUPs were assessed as total rMUP load on column increased. Error bars represent SD (n = 2).

5.2.3 ESI-MS linearity of rMUPs using the Waters Synapt G2

Recombinant forms of MUPs 7, 11 and 20 (darcin) with an N terminal hexa-histidine tag were expressed by Mrs L McLean as described in Chapter 2, and were purified using affinity chromatography, also described in Chapter 2. The concentrations of the purified proteins were confirmed by protein assay, as described previously. To assess the relationship between rMUP concentration and ESI-MS peak area in mixtures and as a single protein, and therefore determine the responses of each of the rMUPs and the factors that affect them, each rMUP was diluted to the same concentration of 3 mg/ml prior to further analysis.

To confirm that all three rMUP samples were at approximately the same concentration, a second protein assay was carried out alongside SDS-PAGE analysis. rMUP samples (all at 3 mg/ml) were diluted ten-fold prior to analysis on a 15% SDS-PAGE gel. The intensities of the three stained protein bands corresponding to each rMUP were similar, indicating that the rMUP concentrations were approximately the same (Figure 5.9). For the protein assay, rMUP samples were diluted 1 in 150. The protein assay confirmed that the three rMUP samples were of very similar concentration (2.8 mg/ml – 3.0 mg/ml) (Figure 5.9), meaning that these samples were suitable for use in the comparison of ESI-MS linearity and response factors of rMUPs.

All future experiments took place using the Waters Synapt G2 Q-ToF mass spectrometer due to the Waters Synapt G1 no longer being available. To determine the dynamic range of this instrument, a number of dilutions of the rMUP 11 sample were performed and analysed (data not shown). A 2 ng protein load presented no multiply charged ion series signal, and therefore the protein could not be detected at this load. A 5 ng load presented a reasonably intense ion series and resultant processed mass spectrum, and a 10 ng protein load resulted in a high intensity multiply charged ion series and true mass spectrum. At a 15 ng protein load, the ESI-MS response was less intense than that at 10 ng, indicating that this protein load resulted in the mass spectrometer detector becoming saturated. A sample of rMUP 11 (5 ng/µl) was injected at a range of injection volumes to result in protein loads of 2.5, 5, 7.5, 10 and 12.5 ng to confirm the dynamic range of the instrument. The lower limit of detection was approximately 2.5 ng and the peak area/protein load relationship remained linear up to a load of 10 ng; at 12.5 ng load, the relationship became non-linear due to detector saturation (Supplementary Material).

The three rMUP samples (of equal concentration) were diluted appropriately for ESI-

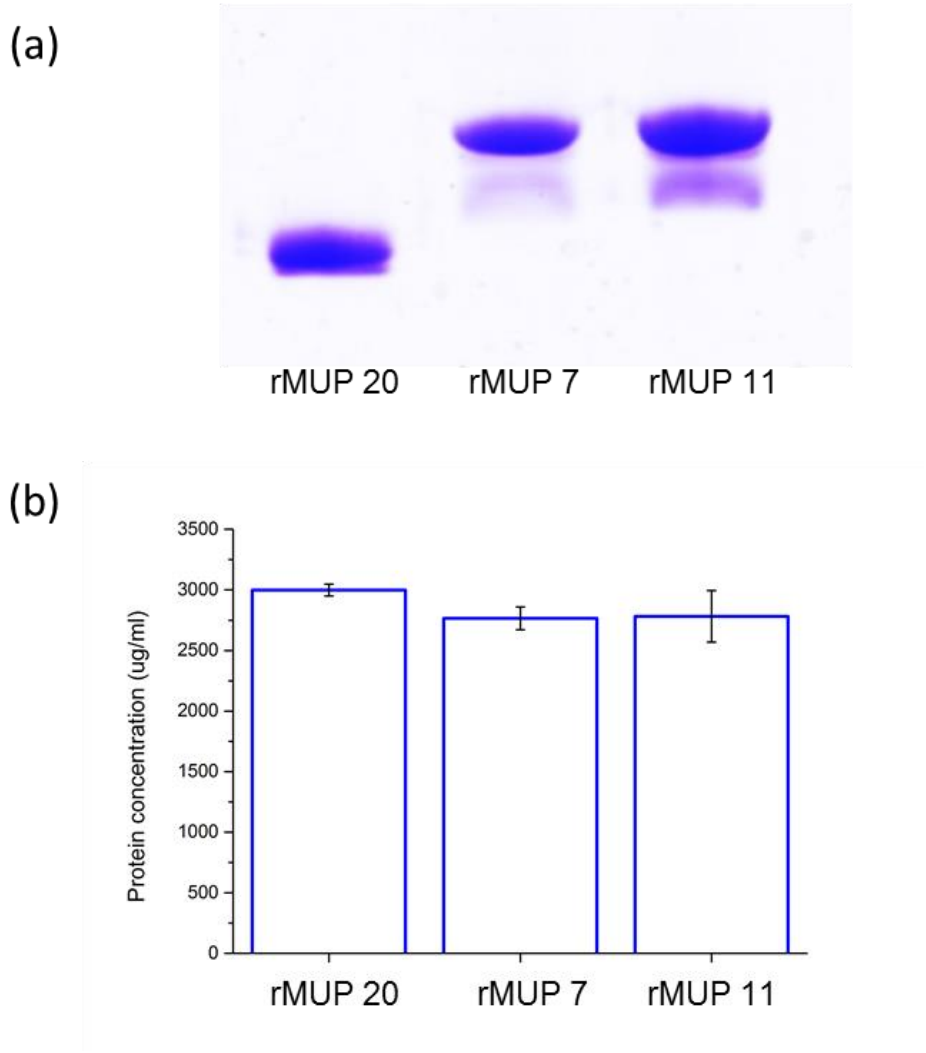


Figure 5.9 a) SDS-PAGE analysis and b) determination of concentration of the four rMUPs.
a) 10 µl of each diluted rMUP sample was mixed 1:1 with sample buffer, loaded onto a 15% (w/v) SDS gel and resolved by SDS-PAGE. The gel was stained with Coomassie blue to visualise the proteins. b) Protein concentration of the four rMUP samples was determined using a Coomassie Plus protein assay kit. A standard curve (0-50 µg/ml) was produced using serum albumin. Samples were diluted 1 in 150 in milliQ water. Absorbance readings were measured at 620 nm and protein concentration was calculated, taking into account dilutions (n = 4, error ± SD).

MS analysis within the dynamic range as confirmed in the analysis of rMUP 11, and each rMUP was analysed separately using ESI-MS to assess the relationship between rMUP concentration and peak area. Each sample was diluted to 5 ng/µl, and a range of injection volumes (0.5 µl – 2 µl) of the samples resulted in total protein loads of 2.5, 5, 7.5 and 10 ng for each rMUP. ESI-MS analysis was carried out using the method described in Chapter 2, and each sample was injected in triplicate. The raw mass spectra were processed as described in Chapter 2, giving the molecular weight of each rMUP and its peak intensity, and the area under peaks were calculated (Figures 5.10 - 5.12). Analysis of the responses of single rMUPs in ESI-MS would determine whether the linearity of ESI-MS response with different single rMUP loads is different for different rMUPs, and if so, whether charge state distribution and/or the number of protonatable sites a single rMUP possesses affects its ionisation, and therefore resulting signal intensity and peak area.

As in the analysis of rMUPs using the Synapt G1 Q-ToF, the charge state distribution profiles in the ESI-MS analysis of rMUPs are different to one another; however, distribution profiles are also different to those observed in the Synapt G1 analysis – profiles in G2 analysis are centred around the $[M+19H]^{19+}$ ion for rMUP 20 (darcin), around the $[M+20H]^{20+}$ ion for rMUP 11 and around the $[M+21H]^{21+}$ ion for rMUP 7 (Figures 5.6 – 5.8). Again, as in the previous experiment using the Synapt G1, a second distribution profile is observed for each rMUP, indicative of the proteins retaining structure under these experimental conditions. In the Synapt G1 analysis of rMUPs, the intensity of the second profile, centred around the $[M+11H]^{11+}$ ion for all rMUPs, was much higher intensity in relation to the 'main' distribution profile than in the Synapt G2 analysis, where the second profiles were centred around the $[M+11H]^{11+}$ ion for rMUPs 7 and 11 but around the $[M+10H]^{10+}$ ion for rMUP 20 (Figure 5.13). This indicates that in the Synapt G1 analyses, where ionisation conditions are the same as in the G2 analysis, more protein is somehow retaining a more folded structure during ESI. However, in the G2 analysis, rMUP 20, with both distribution profiles shifting to a slightly lower charge state distribution than in G1 analyses, appears to remain more folded than rMUPs 7 and 11. When the charge state distribution profiles for all three rMUPs were normalised to 1 and compared (Figure 5.14), it was confirmed that the highly homologous central rMUPs 7 and 11 exhibited very similar charge state distributions to each other, with both having two conformational states in the gas phase. The less homologous rMUP 20 exhibited a different, lower charge state distribution to the central rMUPs in both distribution profiles, suggesting that this rMUP is more folded under these conditions and

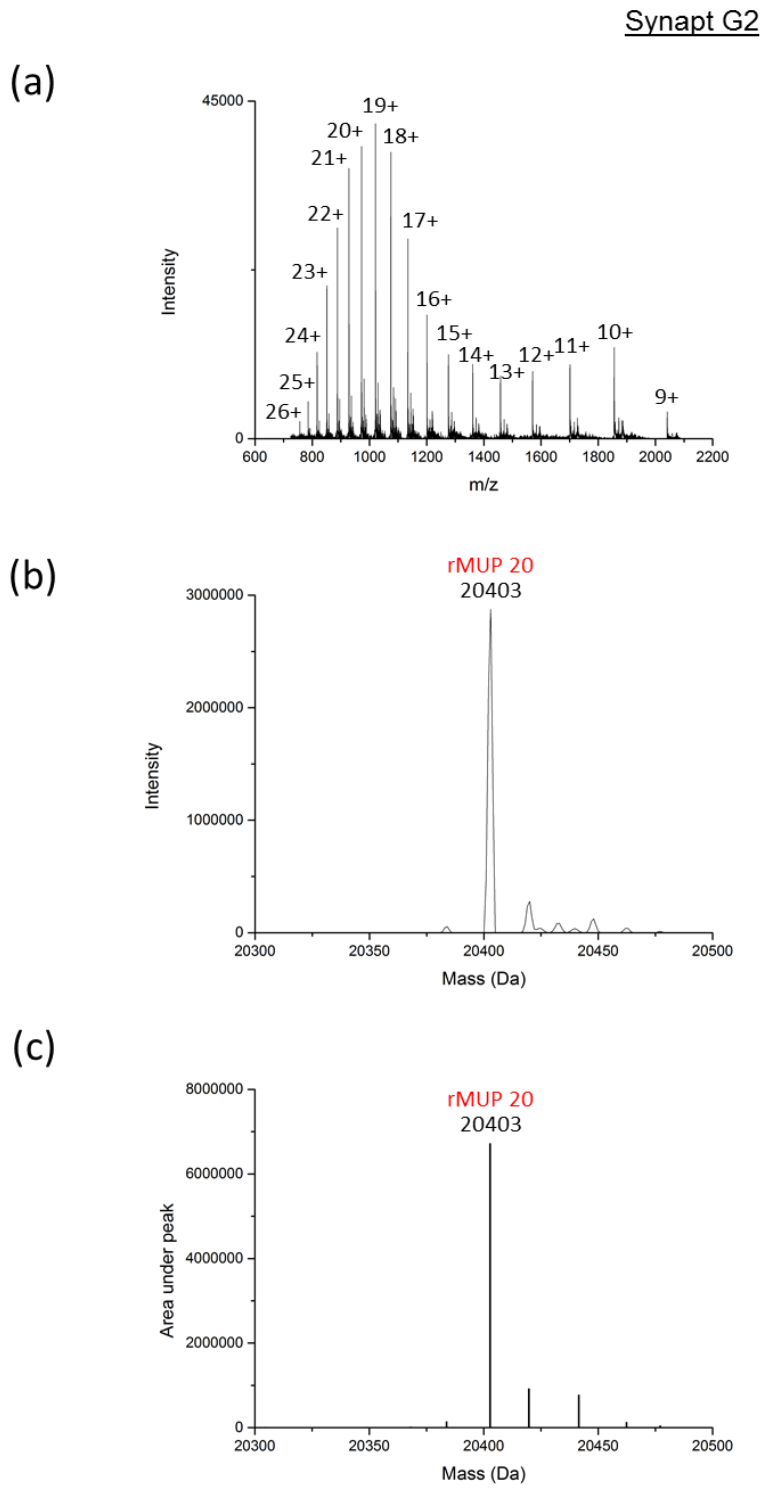


Figure 5.10 ESI-MS analysis of recombinant MUP 20 (Darcin).
rMUP 20 (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap and the mass of rMUP 20 was confirmed by ESI-MS in each case. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). This figure shows the data from the 7.5 ng protein load. a) The multiply charged ion series for rMUP 20. b) The deconvoluted, true mass spectrum, showing the intact mass for rMUP 20. c) The true mass spectrum viewed as a peak area, which gives the relative concentration of the rMUP in each sample injection.

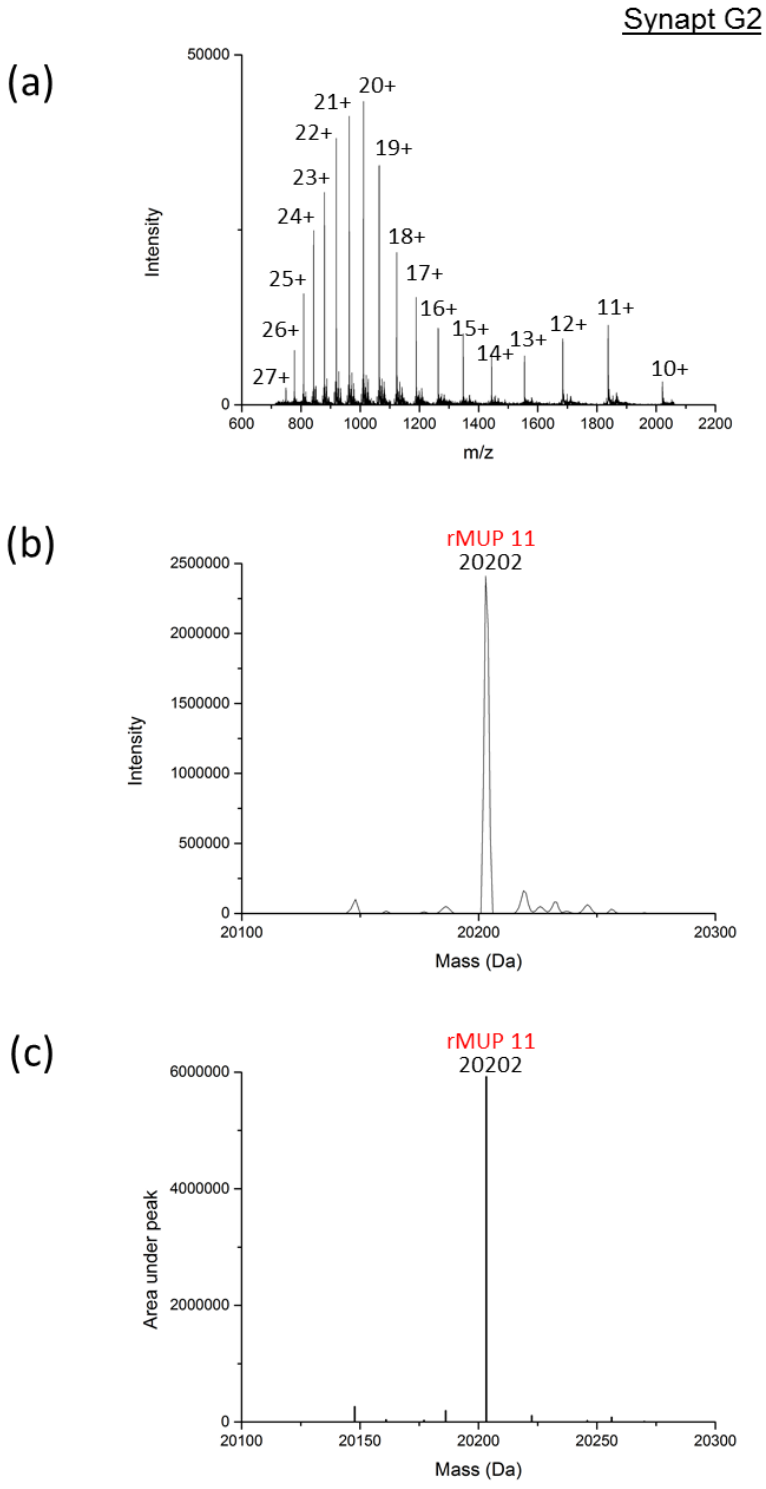


Figure 5.11 ESI-MS analysis of recombinant MUP 11.
rMUP 11 (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap and the mass of rMUP 11 was confirmed by ESI-MS in each case. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). This figure shows the data from the 7.5 ng protein load. a) The multiply charged ion series for rMUP 11. b) The deconvoluted, true mass spectrum, showing the intact mass for rMUP 11. c) The true mass spectrum viewed as a peak area, which gives the relative concentration of the rMUP in each sample injection.

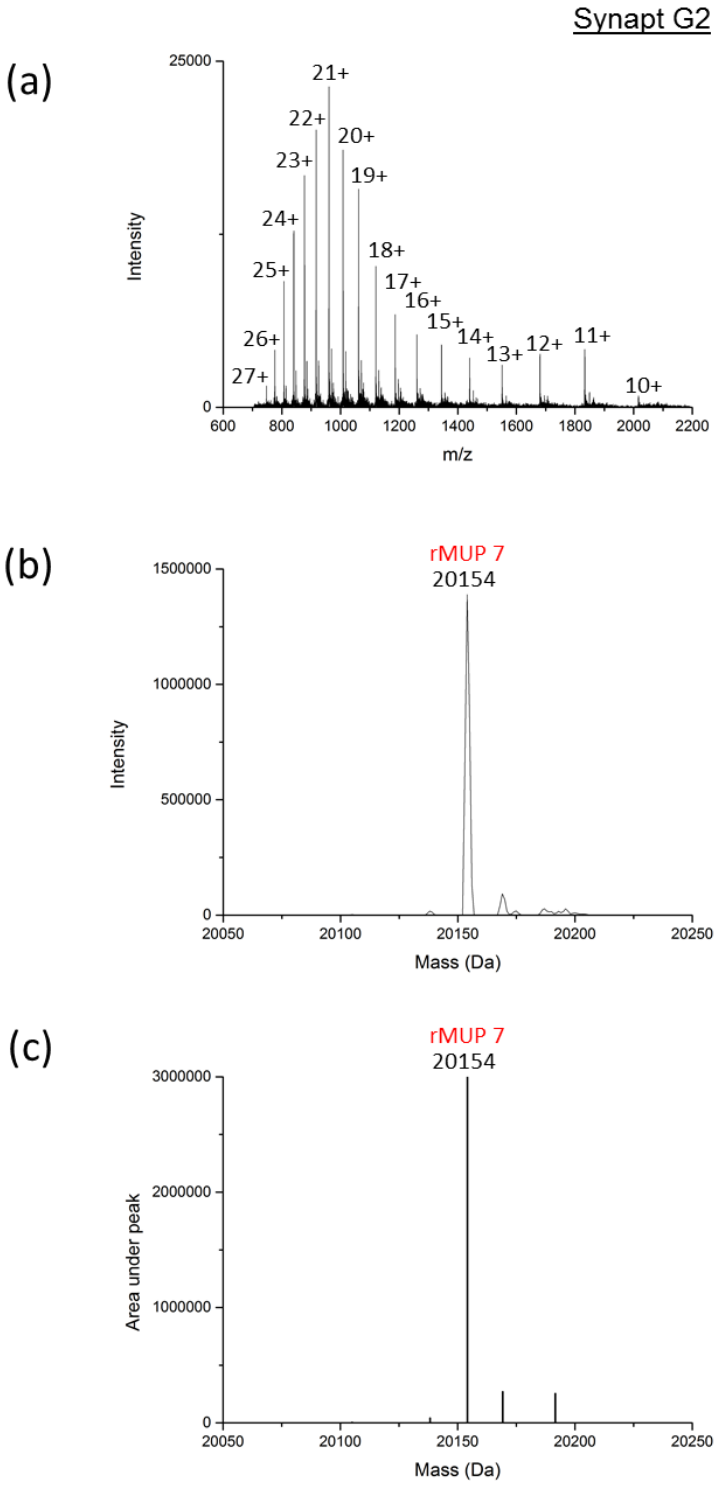


Figure 5.12 ESI-MS analysis of recombinant MUP 7.
rMUP 7 (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap and the mass of rMUP 11 was confirmed by ESI-MS in each case. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). This figure shows the data from the 7.5 ng protein load. a) The multiply charged ion series for rMUP 7. b) The deconvoluted, true mass spectrum, showing the intact mass for rMUP 7. c) The true mass spectrum viewed as a peak area, which gives the relative concentration of the rMUP in each sample injection.

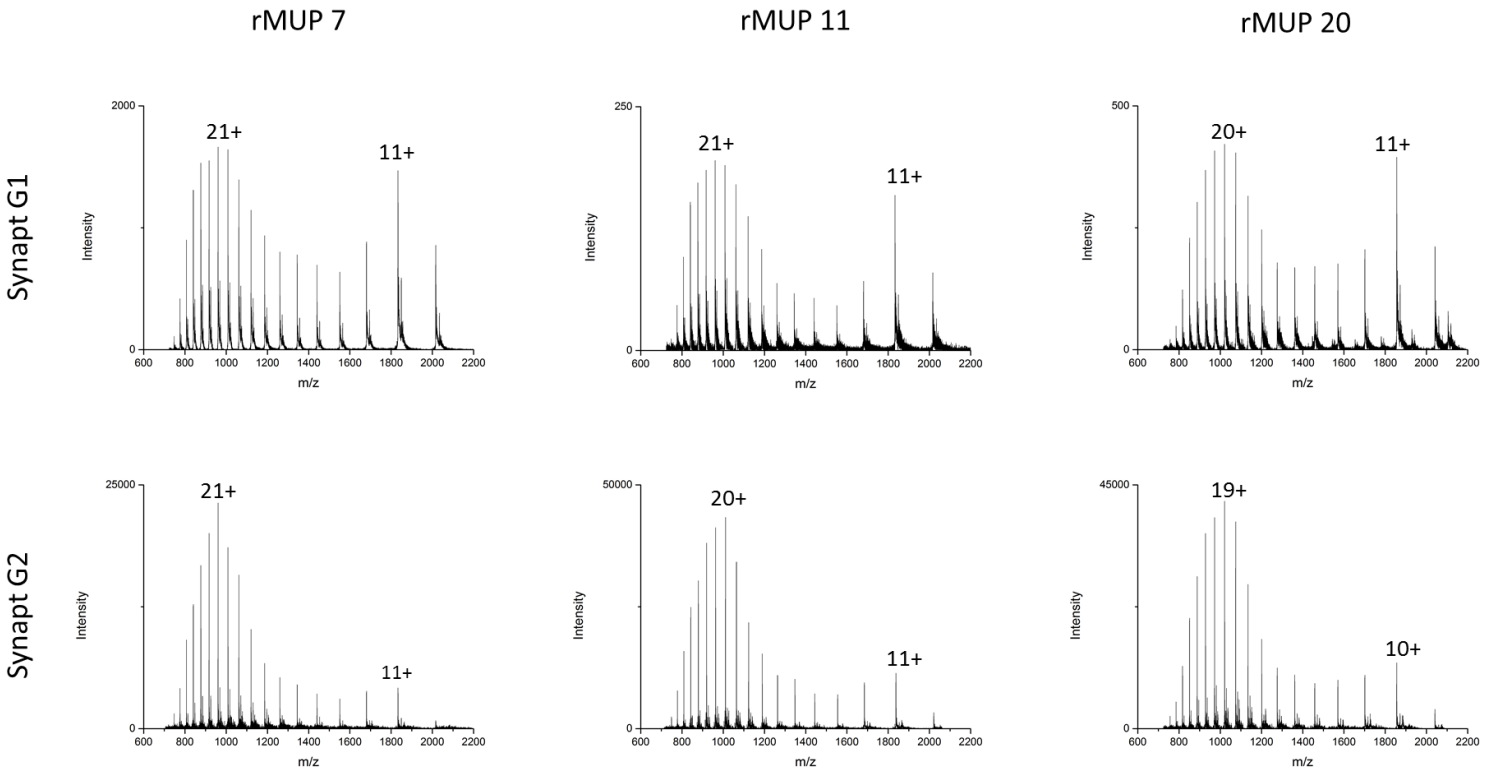


Figure 5.13 Comparison of rMUP charge state distributions in Synapt G1 and Synapt G2 analysis.
The charge distribution profiles observed in Synapt G1 analysis were compared with those observed in Synapt G2 analysis for each rMUP. In Synapt G1 analysis, rMUP 7 displayed two similar-intensity profiles, one centred around the $[M+21H]^{21+}$ ion and one centred around the $[M+11H]^{11+}$ ion. In G2 analysis, the same profiles were observed but the profile around the $[M+11H]^{11+}$ ion was much lower intensity. In the G1 analysis of rMUP 11, the two profiles (around the $[M+21H]^{21+}$ and $[M+11H]^{11+}$ ions) were similar intensity, but in G2 analysis the 'main' profile was centred around the $[M+20H]^{20+}$ ion and the profile around the $[M+11H]^{11+}$ ion was much lower in intensity. For rMUP 20, the G1 analysis showed two similar intensity profiles centred around the $[M+20H]^{20+}$ and $[M+11H]^{11+}$ ions, but G2 analysis showed the distribution profiles centred on the $[M+19H]^{19+}$ and $[M+10H]^{10+}$ ions, with the $[M+10H]^{10+}$ profile much lower in intensity.

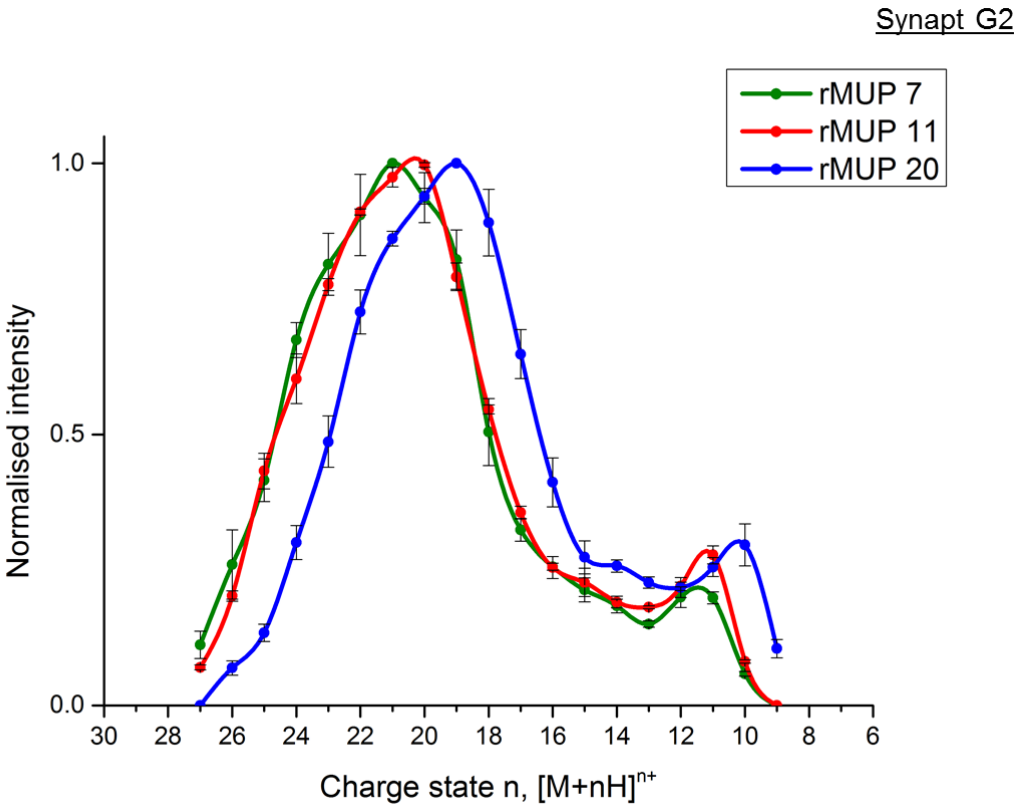


Figure 5.14 Differences in charge state distribution of rMUPs in ESI-MS using the Synapt G2.
The multiply charged ion series for each of the rMUPs in the linearity experiment were observed for differences. The intensities of each of the charge states observed in the raw spectra for each rMUP were recorded and normalised. Error bars represent SD (n=3).

therefore experiencing a lower degree of charging than rMUPs 7 and 11 in both conformational states. This could be explained by the structural analysis of native MUPs 11 and 20, where Phelan *et al*. (2014) were able to confirm that MUP 20 is significantly more structurally stable than MUP 11, resulting in part by the orientation of the disulphide bond (Phelan *et al.* 2014). To determine whether the differences in protonation of each rMUP in ESI-MS analysis (resulting from the lower degree of charging of rMUP 20 than rMUPs 11 and 7 due to its more compact structure in the gas phase) led to variability in the intensity of ions generated by each rMUP, the responses of each rMUP at different protein loads were recorded and compared.

The responses of the peak area values with protein load in ESI-MS analysis for rMUPs 7, 11 and 20 were assessed (Figure 5.15) and compared (Figure 5.16). rMUPs 7 and 11 exhibited a linear relationship between protein load and peak area, with peak areas increasing as protein load increased, and minimal variation in calculated peak areas amongst technical replicates (Figures 5.15 and 5.16) These rMUPs exhibited tight linear relationships, with r^2 values of 0.99. The relationship between peak area and protein load for rMUP 20 was non-linear ($r^2 = 0.99$), where the peak area values increase disproportionately with increasing load of rMUP. It was expected that rMUP 20, with the lower charge state distribution, would have a linear response in ESI-MS analysis that was a shallower slope than those for rMUPs 7 and 11 due to fewer accessible protonatable sites, poorer ionisation and thus, lower signal intensity in the mass spectra. For central rMUPs 7 and 11, it was expected that the linear slopes would be very similar, due to the similar charge state distributions rising from a similar number of protonatable sites. Figure 5.15 gave a numerical comparison of the linear peak area/protein load responses by rMUPs 7 and 11 in ESI-MS analysis, whilst Figure 5.16 provided a visual comparison. Unexpectedly, for rMUP 7, the gradient of the slope was considerably shallower than that of rMUP 11, meaning that a lower ESI-MS response per ng was observed compared to rMUP 11. The non-linear power function fitted to the peak area/protein load relationship exhibited by rMUP 20 meant that at a 2.5 ng rMUP load, the peak area was approximately the same as that for rMUP 7, lower than that for rMUP 11 at 5 ng load but significantly greater than both rMUPs 7 and 11 at 7.5 ng and 10 ng loads.

Whilst the peak areas for each rMUP, obtained through deconvolution of the multiply charged protein envelopes, increased with increasing protein load, the peak areas for rMUP 20 increased disproportionally and the peak areas for rMUPs 7 and 11 increased proportionally. The difference in peak areas between the different rMUPs

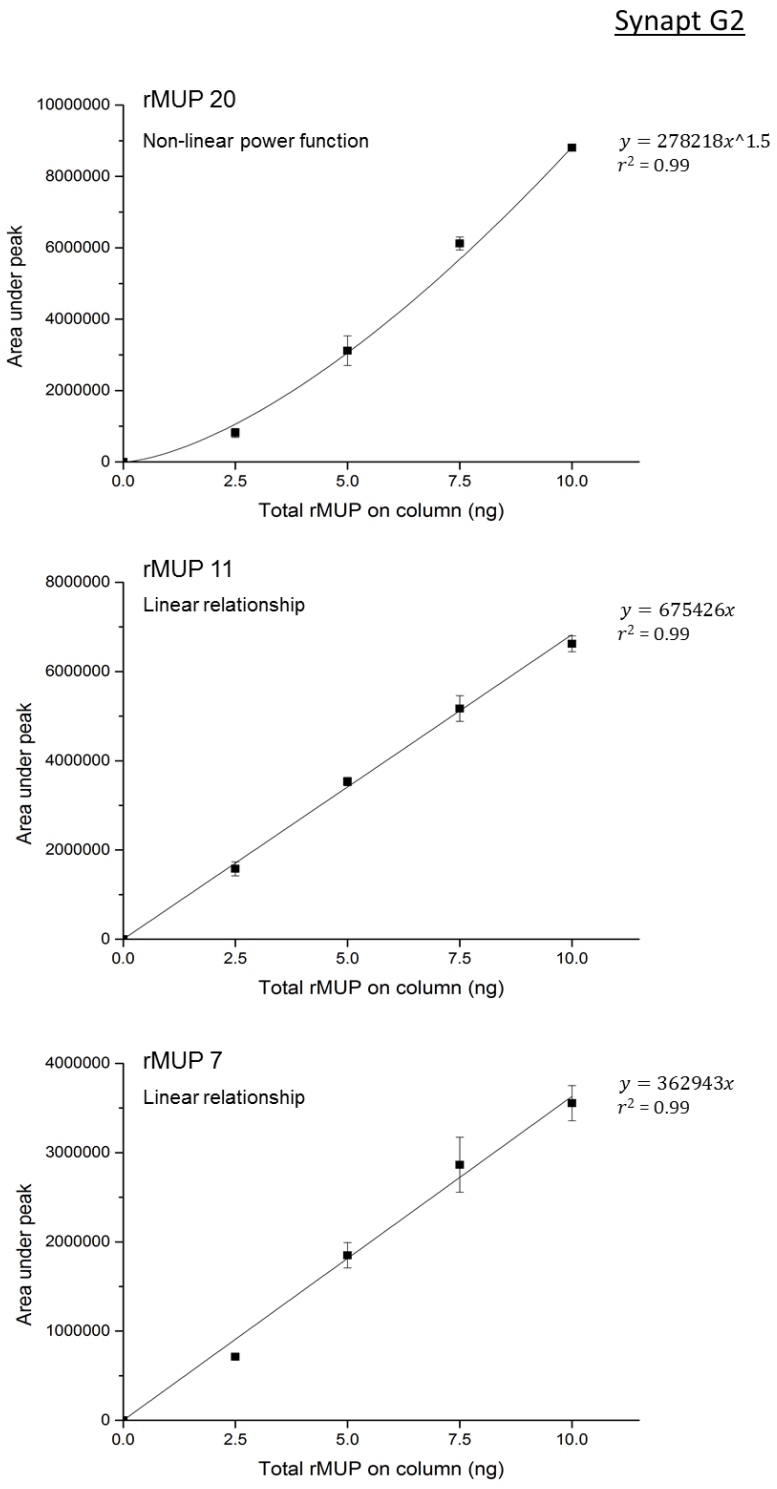


Figure 5.15 Relationship of ESI-MS peak area with increasing protein load for three rMUPs.
Samples of rMUPs 20, 11 and 7 (diluted to 5 ng/µl in 0.1% (v/v) formic acid) were injected over a range of injection volumes (0.5 - 2 µl), resulting in a range of protein loads, onto a C4 desalting trap. The mass of each rMUP was confirmed by ESI-MS. Each sample analysis was in triplicate. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). The area under peak values for each rMUP were assessed as total protein load on column increased. Error bars represent SD (n=3).

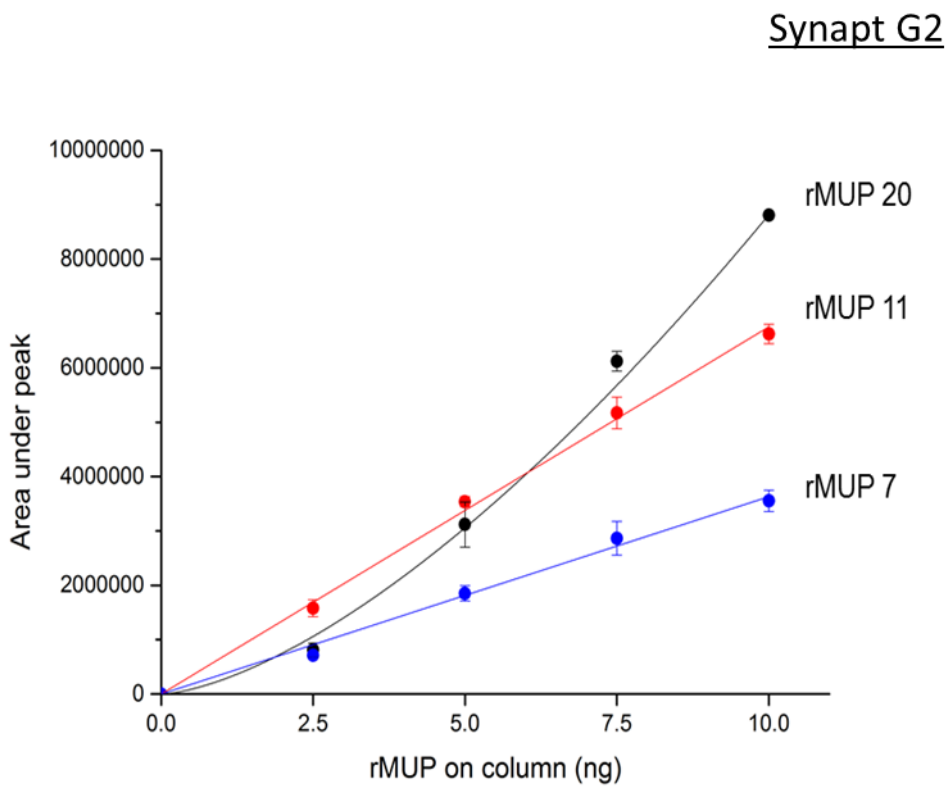


Figure 5.16 Differences in ESI-MS peak area/protein load responses for three rMUPs. The area under peak values for each rMUP were assessed as total protein load on column increased in Figure 5.15. These were plotted on the same scale to visualise the differences in the relationship of ESI-MS peak area with rMUP load between each rMUP. Error bars represent SD (n=3).

at the same protein loads indicates that intact rMUPs have different responses and ionisation efficiencies in ESI-MS analysis. These differences cannot be explained by different charge state distributions or the number of protonatable sites the protein possesses. However, the responses exhibited by each rMUP, although different to one another, are similar enough to confirm that intact mass analysis of MUPs may be of use for the relative quantification of MUPs in a sample (Figure 5.16).

To determine whether ESI-MS of intact proteins is suitable for absolute quantification of MUPs in mouse urine, the next experiment needed to confirm that the three rMUPs would exhibit the same responses in ESI-MS analysis whilst in an equimolar mixture. If this was the case, the ESI-MS responses of rMUPs (and therefore natural MUPs in urine samples) could be determined, meaning that rMUPs corresponding to the natural MUPs in a urine sample could be added to urine samples in a known quantity to produce standard curves for absolute quantification, with rMUP and MUP responses not being affected by sample complexity. The use of an rMUP for each MUP in the sample would account for the differences in ionisation efficiencies observed in this experiment.

5.2.4 ESI-MS responses of rMUPs in an equimolar mixture

The purified rMUP 7, 11 and 20 samples (of equal concentration) used for the analysis of the ESI-MS linearity of single rMUPs were mixed 1:1:1 and diluted to 5 ng/µl. Protein assays were undertaken to confirm that the all samples were still of equal concentration prior to mixing. A range of injection volumes (0.5 µl – 2 µl) of the samples resulted in total protein loads of 2.5, 5, 7.5 and 10 ng for the equimolar rMUP mixture. Therefore, at each injection, 0.8, 1.7, 2.5 and 3.3 ng of each rMUP was loaded on column. It was not possible to load 2.5, 5, 7.5 and 10 ng of each rMUP (in an equimolar mixture) as this would have resulted in total protein loads above the dynamic range of the instrument, resulting in detector saturation. The peak areas for each rMUP at each total protein load would be expected to be smaller than those recorded for the single rMUP samples (since each rMUP in the equimolar mixture accounts for only one third of the total protein load), but the rMUP responses in ESI-MS analysis would be expected (and required) to be the same if intact protein analysis was suitable for the absolute quantification of MUPs. Based on the peak area/protein load ESI-MS responses observed for the single samples of rMUPs 7, 11 and 20 (Figure 5.15), theoretically, the three-sample equimolar mixture would result in each rMUP having a peak area of approximately a third of the value it exhibited in a single rMUP sample at each protein load, and the gradient of the slope would be expected

to be approximately a third of the value. For rMUP 20, where the ESI-MS response was a non-linear power function (y = ax^b), the coefficient parameter 'a' (that specifies slope) would be expected to have a value of approximately a third of the value calculated for the single-rMUP sample, and the power parameter 'b' (that specifies the rate at which peak area disproportionally increases with protein load) should have the same value.

ESI-MS analysis and data processing was as described in Chapter 2, giving the molecular weight of each rMUP and its peak intensity, and the area under peaks (and therefore relative rMUP concentration) were calculated in the equimolar mixture. Each sample was injected in triplicate. The charge state distribution profiles of the three rMUPs in the mixture were studied, and the profiles in the 7.5 ng equimolar mixture are shown in Figure 5.17. Inset in the Figure is an enlargement of one of the charge states, showing the rMUP 7 $[M+18H]^{18+}$ ion which is significantly lower in intensity than rMUPs 11 and 20, concurring with the ESI-MS response exhibited by this rMUP in a single-rMUP sample (Figures 5.15 and 5.16) and the charge state distribution profile observed for this protein (Figure 5.14). At a 7.5 ng protein load, based on the similar responses confirmed for the single-rMUP samples for rMUPs 11 and 20 (Figure 5.16) and the charge state distribution profiles for these rMUPs (Figure 5.14), it would be expected that the rMUP 11 $[M+26H]^{26+}$ - $[M+20H]^{20+}$ ions would be higher intensity than the corresponding MUP 20 ions, and that the $[M+19H]^{19+}$ - $[M+12H]^{12+}$ ions would be lower intensity. Whilst the lower charge states of rMUP 11 were lower intensity than rMUP 20 as expected, rMUP 20 was more intense than rMUP 11 in the mass spectra at the $[M+23H]^{23+}$, $[M+22H]^{22+}$ and $[M+21H]^{21+}$ ions (Figure 5.17). This was the case in the spectra of 7.5 and 10 ng protein loads, but at 2.5 and 5 ng loads, the intensity of the rMUP 11 $[M+26H]^{26+}$ - $[M+20H]^{20+}$ ions were higher intensity than the corresponding rMUP 20 ions, as they were in single-rMUP sample analyses (Figure 5.17). This indicates that at higher loads of the three-rMUP equimolar mixture, rMUP 20 exhibits a different multiply charged ion series than in single rMUP sample ESI-MS analysis, occupying a higher charge state distribution. The mass spectra were deconvoluted using the entire multiply charged ion series to ensure a true representation of the peak area/protein load ESI-MS responses of each rMUP in the equimolar mixture – if only part of the spectrum was analysed for molecular weight and peak area information, protein molecular weights would be accurate but peak areas calculated by the MaxENT software would be likely to be inaccurate. For example, if spectra were only processed across the $[M+19H]^{19+}$ - $[M+12H]^{12+}$ ion range (where rMUP 20 ions are higher intensity than rMUP 7 and 11

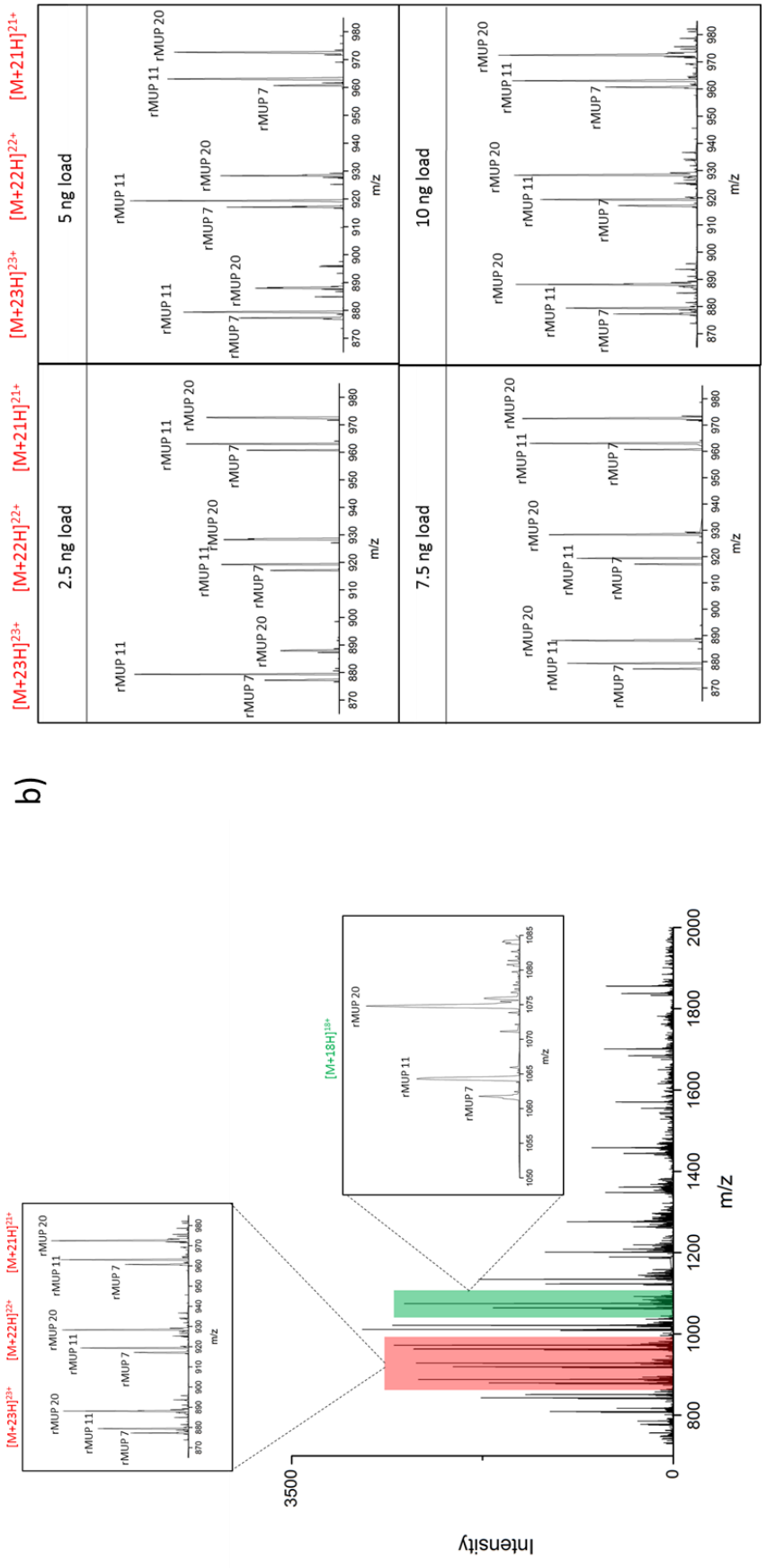


Figure 5.17 Charge state distributions of rMUPs 7, 11 and 20 in an equimolar mixture.
The equimolar rMUP mixture (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap. (a) The multiply charged ion series observed in the 7.5 ng load. Inset is an enlargement of the $[M+18H]^{18+}$ ions (green) and an enlargement of the $[M+23H]^{23+}$,$[M+22H]^{22+}$and $[M+21H]^{21+}$ ions (red). (b) The $[M+23H]^{23+}$,$[M+22H]^{22+}$and $[M+21H]^{21+}$ ion intensities for each rMUP in the 2.5, 5, 7.5 and 10 ng loads.

ions), this would result in larger rMUP 20 and smaller rMUP 7 and 11 peak areas being calculated than if the entire ion series was used to calculate the peak areas of each rMUP. The ESI-MS intact protein peak areas of rMUPs 7, 11 and 20 in the equimolar mixture at 2.5, 5, 7.5 and 10 ng loads (Figure 5.18) were plotted (Figure 5.19) and the responses of the peak area values with protein load for each of the rMUPs were compared with each other and with the theoretical ESI-MS responses via the gradient of the slopes. Figure 5.18 shows the deconvoluted true mass spectra and corresponding peak area spectra of the equimolar rMUP mixture at each protein load. At 2.5 ng total protein load, the calculated peak areas for rMUPs 11 and 20 were similar, with the rMUP 11 peak area being slightly greater. The peak area for rMUP 7 is around 50 % of the peak area values obtained for the other two rMUPs in the equimolar mixture. As protein load increased, the peak areas calculated for rMUP 20 increased significantly relative to those calculated for rMUPs 7 and 11, with the peak areas calculated for these two rMUPs decreasing relative to rMUP 20 (Figure 5.18). This initial analysis suggests that in an equimolar rMUP mixture, the peak area/protein load relationship exhibited by rMUP 20 is a non-linear power function, as per the analysis of this rMUP as a single protein (Figure 5.15). To confirm whether the three rMUPs exhibited the same responses in ESI-MS analysis whilst in an equimolar mixture to those observed in the single rMUP analyses, the responses of the peak area values with protein load in ESI-MS analysis for rMUPs 7, 11 and 20 were assessed and compared. In the equimolar mixture, rMUP 20 exhibited a non-linear relationship between peak area and protein load (r^2 = 0.98), as it did in the single rMUP analysis. Whilst rMUPs 7 and 11 exhibited linear peak area/protein load relationships in the single rMUP analyses, as part of an equimolar mixture, both rMUPs exhibited a non-linear relationship (r^2 = 0.99), meaning peak area values increase disproportionately with increasing load of rMUP (Figure 5.19). Comparing the power parameters of the non-linear power function fits for each rMUP, the responses of rMUPs 11 and 20 increase at the same rate (with a ‘b’ value of 1.8), but a higher ‘b’ value of 2.3 fitted to the rMUP 7 slope indicates that the rMUP 7 peak areas increase with protein load at a greater rate. The analysis of the peak area/protein load responses of rMUPs 7 and 11 confirm that they do not exhibit the same response in ESI-MS analysis as part of a mixture as they do in a single-rMUP sample (Figure 5.19). To determine whether rMUP 20 exhibited the same response as part of a mixture as in a single-rMUP sample, the ESI-MS response in the equimolar mixture was plotted with the single-rMUP response and the theoretical response for rMUP 20 in the three-rMUP mixture (Figure 5.20). The gradient of the

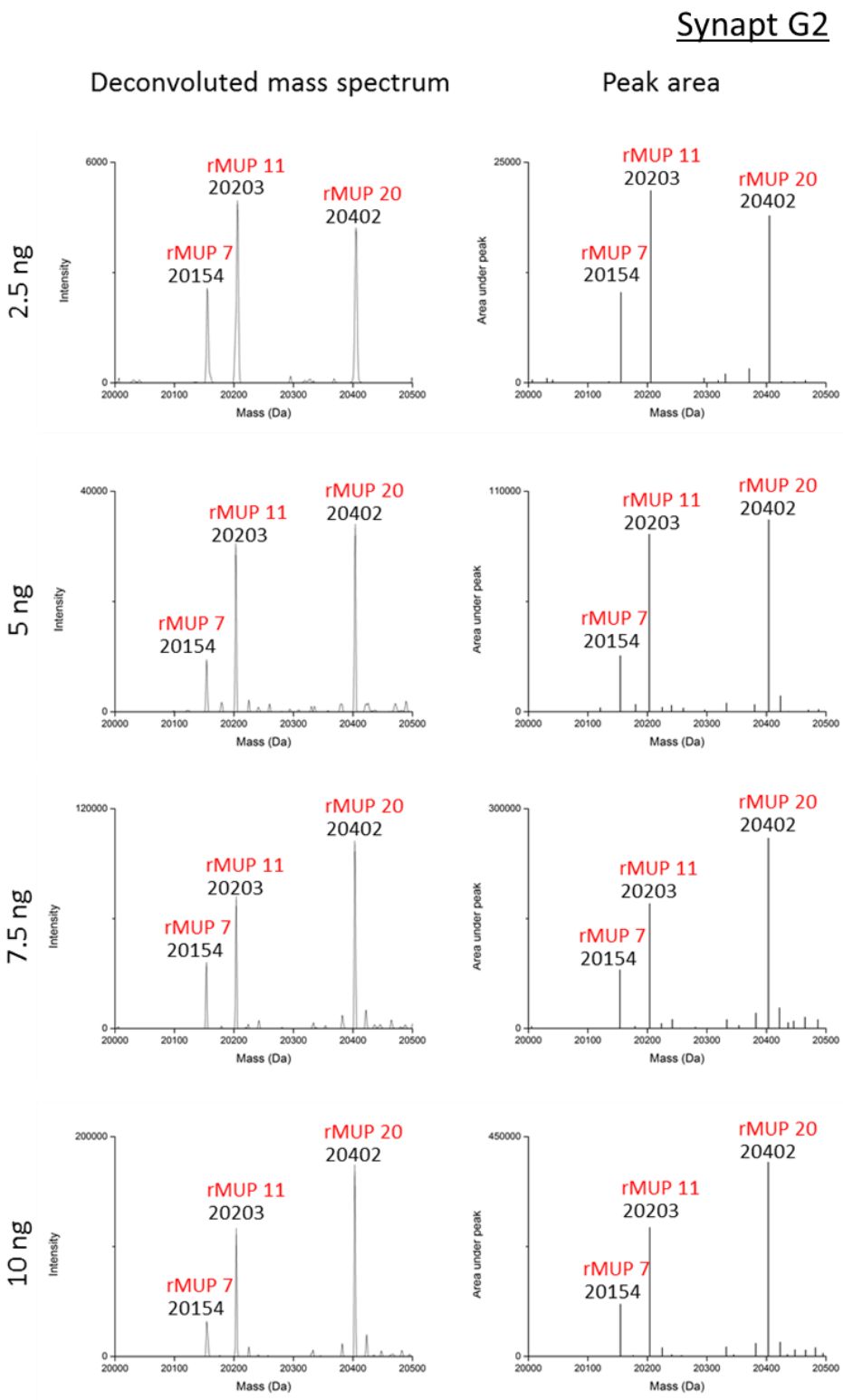


Figure 5.18 ESI-MS analysis of the three-rMUP equimolar mixture at 2.5, 5, 7.5 and 10 ng total protein loads.
The equimolar mixture of rMUPs 7, 11 and 20 (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap and the masses of the rMUPs were confirmed by ESI-MS in each case. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). The figure shows the deconvoluted, true mass spectra, showing the intact masses for the rMUPs, and the true mass spectra viewed as a peak area.

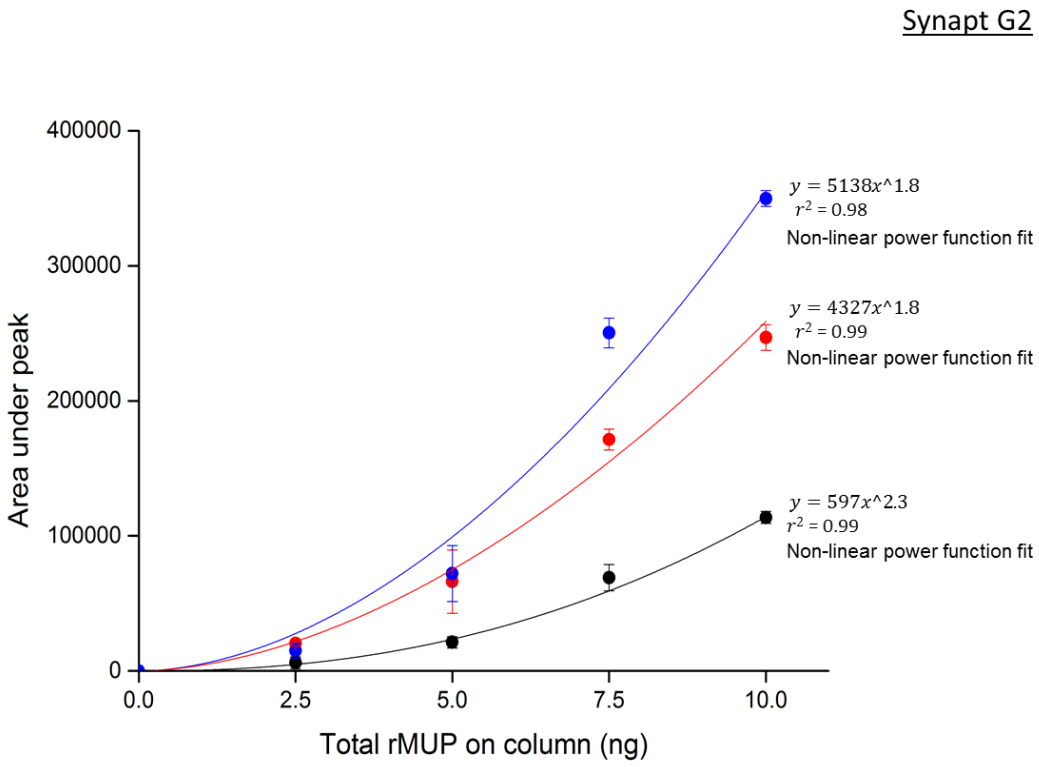


Figure 5.19 Differences in peak area with increasing protein load for the three rMUPs in the equimolar mixture.
The area under peak values for each rMUP were assessed as total protein load on column increased. These were plotted on the same scale to visualise the differences in the ESI-MS responses between each rMUP as part of an equimolar mixture. Error bars represent SD (n=3).

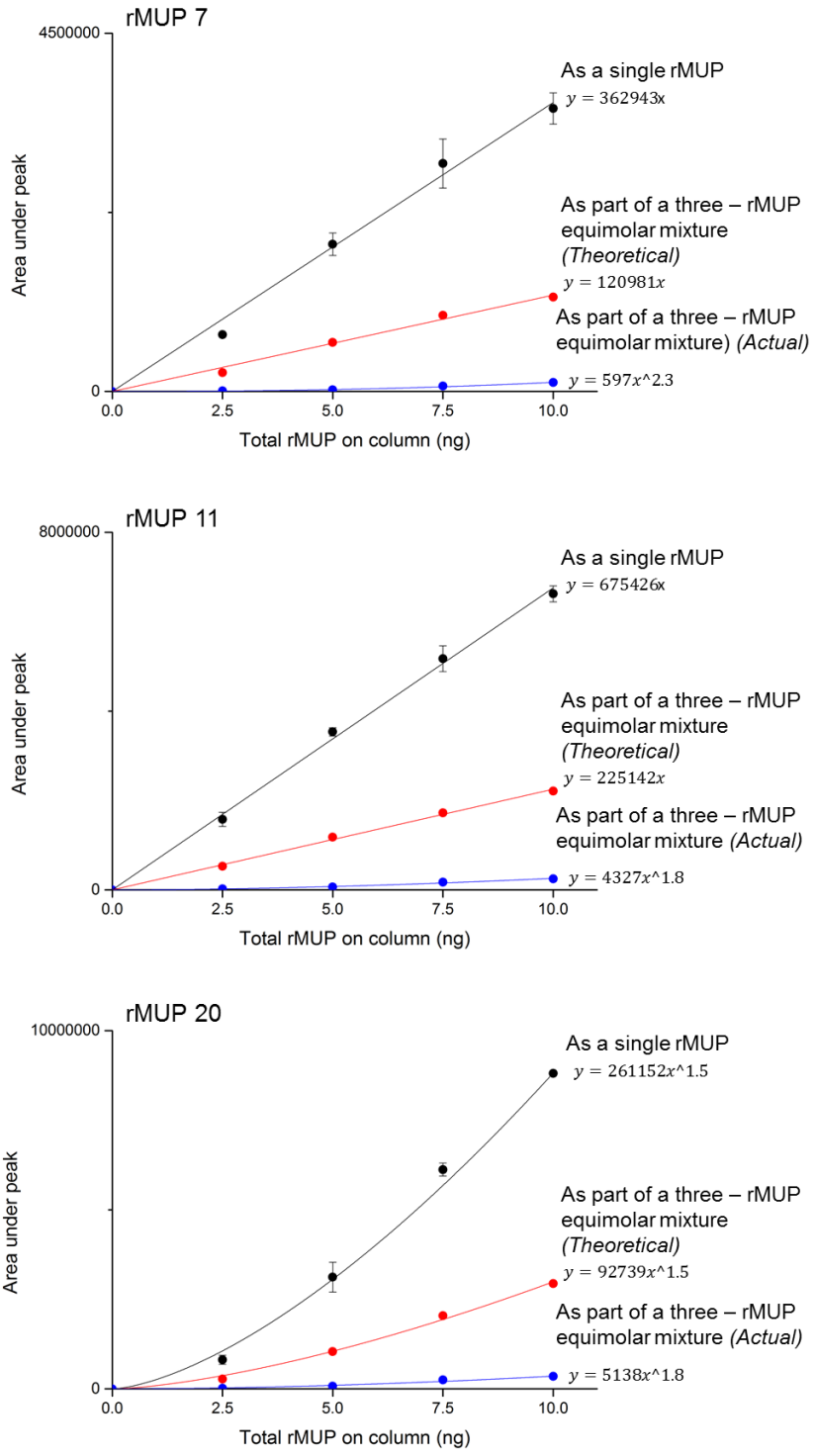


Figure 5.20 Comparing the different ESI-MS responses exhibited by each rMUP as part of a three-rMUP equimolar mixture and as a single rMUP sample.
The area under peak values for each rMUP as total protein load on column increased were recorded for the proteins as single samples (black data points) as well as part of a three-rMUP equimolar mixture (blue data points). These were plotted on the same scale, along with theoretical area under peak values expected for the rMUPs in the three-rMUP mixture if the rMUPs were to exhibit the same ESI-MS responses as in single rMUP samples (red data points). This enabled the visualisation of the differences in the theoretical and actual ESI-MS responses exhibited by each rMUP as part of an equimolar mixture. Error bars represent SD (n=3).

non-linear slope for the observed peak area/protein load response for rMUP 20 in the equimolar mixture was significantly shallower than the theoretical slope (a coefficient value approximately 18 times smaller than expected), but the rate at which peak area disproportionally increased with protein load was higher than expected (a power value of 1.8, in comparison to the value of 1.5 expected from the single rMUP analysis). This confirms that along with rMUPs 7 and 11, rMUP 20 does not exhibit the same response in ESI-MS analysis as part of an equimolar mixture as it does in a single-rMUP sample.

The significant differences in ESI-MS responses exhibited by all three rMUPs when comparing the analysis of single-rMUP samples with equimolar mixtures indicate that ESI-MS of intact proteins is not currently suitable for the absolute quantification of MUPs in mouse urine. This experiment was repeated numerous times to ensure errors had not been made with respect to protein concentration calculations or sample dilutions, so there was confidence in the unusual results obtained for the equimolar mixtures. rMUPs, or any isotopically labelled MUP protein standard, would not be suitable internal standards for absolute quantification since the response factors of rMUPs (and therefore MUPs) are different depending on the complexity of the sample. Therefore, the ESI-MS responses of rMUPs and MUPs cannot be definitively determined in mouse urine samples where complexity varies (e.g. female C57BL/6 urine contains fewer MUPs than male C57BL/6 urine, and MUP expression patterns can vary significantly amongst wild mice (Beynon *et al*. 2002; Cheetham *et al*. 2008)). Despite this, it was noted that in both single-rMUP and equimolar mixture cases, at the 2.5 ng total protein load, the peak area of the three rMUPs were fairly similar to one another; at 5 ng, the peak areas of rMUPs 11 and 20 were similar and rMUP 7 lower; and at 7.5 ng and 10 ng, rMUP 20 peak area was slightly higher than rMUP 11, and rMUP 7 lower than rMUP 11. When comparing the differences in responses observed in each rMUP (Figure 5.14), it was also apparent that all three rMUPs ionised less efficiently to a very similar extent to one another in the equimolar mixture. This could possibly be due to a proton deficit during the ionisation process, where the three rMUPs in the mixture are competing for charge and so all proteins are ionising less successfully than as single-rMUP samples. The next experiment was to determine whether this pattern of ESI-MS responses was also observed in a less complex equimolar rMUP sample; if so, ESI-MS analysis of intact MUPs can still be a suitable method for general relative quantification of MUPs present in urine samples, since the ionisation efficiencies of the three rMUPs have remained the same relative to one another despite different sample complexities. Also, if the ESI-MS

response of the rMUPs as part of a two-sample mixture were not as low as those observed in three-rMUP mixture analysis, but not as expected based on the single-rMUP analysis, this could further indicate a proton deficiency in the ionisation process. Compared to the three-rMUP analysis, a two-rMUP sample would give rise to less competition for charge and therefore more successful protein ionisation, but compared to the single-rMUP analysis, more competition for charge would result in less successful protein ionisation and therefore a lower than expected ESI-MS response.

5.2.5 ESI-MS responses of rMUPs 7 and 11 in an equimolar mixture

The aims of this experiment were to firstly determine whether the same pattern of ESI-MS responses observed in the previous experiments due to differing ionisation efficiencies remained the same in a two-rMUP sample as a single- and three-rMUP sample, confirming whether the ESI-MS analysis is suitable for relative quantification of intact MUPs in urine, and secondly to determine whether a proton deficit in the ionisation process could be the reason for the unexpectedly low ESI-MS responses of rMUPs as part of a three-rMUP mixture. The rMUPs 7 and 11 were chosen for the analysis of a two-rMUP equimolar mixture, and the purified rMUP 7 and 11 samples (of equal concentration) used in previous analyses were mixed 1:1 and diluted to 5 ng/µl. A range of injection volumes (0.5 µl – 2 µl) of the samples resulted in total protein loads of 2.5, 5, 7.5 and 10 ng for the equimolar rMUP mixture (meaning that at each injection, 1.25, 2.5, 3.75 and 5 ng of each rMUP was loaded on column). As in the previous experiment, loading 2.5, 5, 7.5 and 10 ng of each rMUP (in an equimolar mixture) would have resulted in total protein loads above the dynamic range of the instrument, resulting in detector saturation. Theoretically, the peak areas for each rMUP at each total protein load would be expected to be approximately half the value, and the gradient of the slope would be expected to be approximately a half of the value of those observed in the single-rMUP analyses. From the analysis of the three-rMUP equimolar mixture, it was expected that this would not be the case as it was confirmed that the ESI-MS responses of rMUPs cannot be definitively determined.

ESI-MS analysis and data processing was as described in Chapter 2, giving the molecular weight of each rMUP and its peak intensity, and the area under peaks (and therefore relative rMUP concentration) were calculated in the equimolar mixture. Each sample was injected in triplicate. The charge state distribution profiles of the two rMUPs in the mixture were studied, and the profiles in the 7.5 ng equimolar

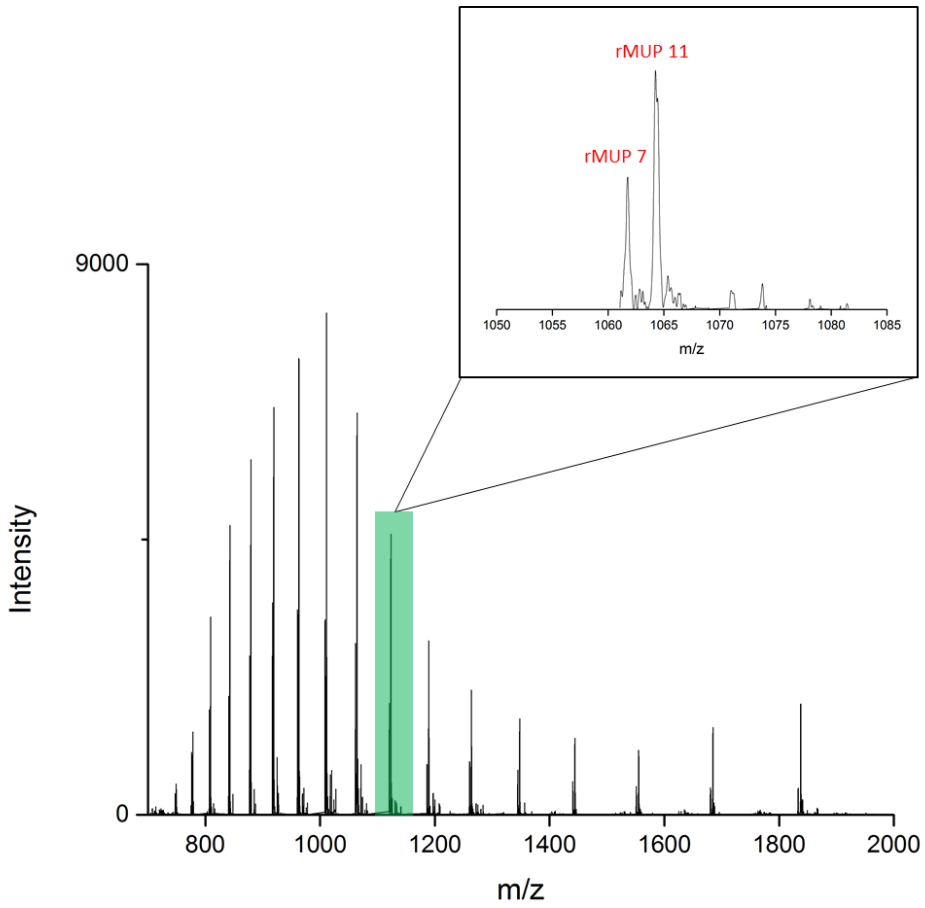


Figure 5.21 Charge state distributions of rMUPs 7and 11 in an equimolar mixture. The equimolar rMUP mixture (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap. The figure shows the multiply charged ion series observed in the 7.5 ng load. Inset is an enlargement of the rMUP 7 and 11 $[M+18H]^{18+}$ ions (highlighted in green).

mixture are shown in Figure 5.21. Inset in the Figure is an enlargement of one of the charge states, showing the rMUP 7 $[M+18H]^{18+}$ ion which is, again, significantly lower in intensity than the rMUP 11 $[M+18H]^{18+}$ ion. The charge state distribution profiles observed in the spectrum in Figure 5.21 are representative of those observed in all protein loads. The multiply charged ion series exhibited by both rMUP 7 and 11 as part of a two-rMUP mixture are similar to those in the analysis of these rMUPs as single samples and in a three-rMUP mixture, suggestive that relative ionisation efficiencies of rMUPs remain the same regardless of sample complexity.

The mass spectra were deconvoluted using the entire multiply charged ion series to ensure a true representation of the peak area/protein load ESI-MS responses of each rMUP in the equimolar mixture. The ESI-MS intact protein peak areas of rMUPs 7 and 11 at 2.5, 5, 7.5 and 10 ng total protein loads (Figure 5.22) were plotted (Figure 5.23) and the responses of the peak area values with protein load for each of the rMUPs were compared with each other and with the theoretical ESI-MS responses via the gradient of the slopes. Figure 5.23 is the deconvoluted true mass spectra and corresponding peak area spectra of the equimolar rMUP mixture at each protein load. At each protein load, the calculated peak areas for rMUP 7 were around 40 – 50 % of the rMUP 11 peak area, as they were for these proteins in the single- and three-rMUP analyses. This is again indicative that the relative ionisation efficiencies of rMUPs are the same regardless of sample complexity. The responses of the peak area values with protein load in ESI-MS analysis for rMUPs 7 and 11 were assessed and compared. In the previous analysis of the three-rMUP equimolar mixture, these rMUPs exhibited non-linear peak area/protein load relationships and linear relationships in the single rMUP analyses; in this experiment, both rMUPs exhibited linear peak area/protein load relationships ($r^2 = 0.99$) (Figure 5.23). The actual ESI-MS responses exhibited by rMUPs 7 and 11 in the two-rMUP equimolar mixture were plotted with the theoretical responses calculated, along with the single-rMUP responses observed (Figure 5.24). As in the three-rMUP equimolar mixtures, rMUPs 7 and 11 exhibited shallower gradients of the linear slopes for the peak area/protein load responses than the theoretical slopes calculated from the single-rMUP analyses. The gradients of the rMUPs 7 and 11 ESI-MS response slopes both had a value approximately 1.7 times lower than the theoretical value.

In the single-rMUP, two-rMUP and three-rMUP sample analysis, the ESI-MS response exhibited by rMUP 7 was lower than those by rMUP 11, as portrayed by the shallower gradients plotted for the peak area/protein load relationships. Taking into

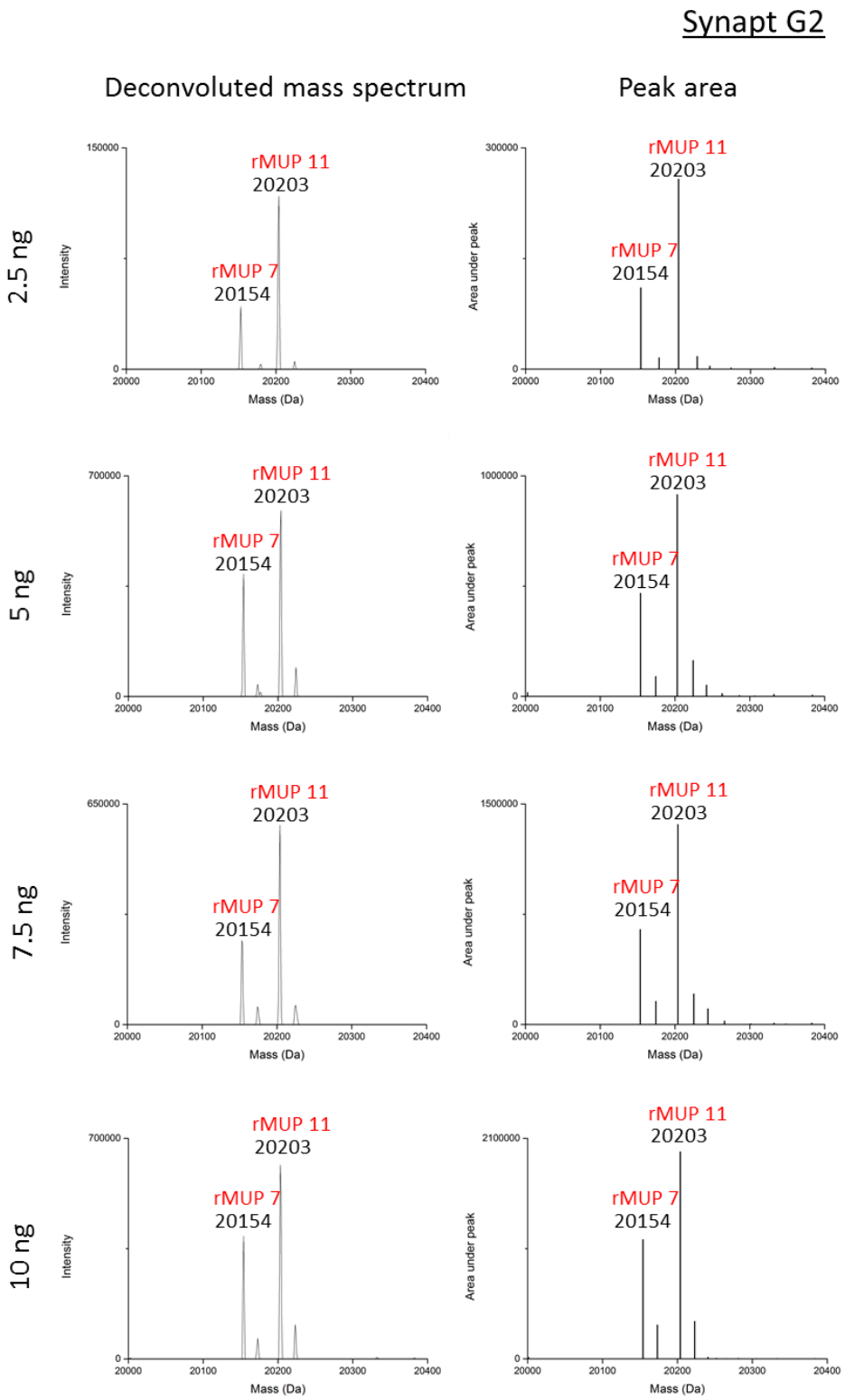


Figure 5.22 ESI-MS analysis of the two-rMUP equimolar mixture at 2.5, 5, 7.5 and 10 ng total protein loads.
The equimolar mixture of rMUPs 7 and 11 (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap and the masses of the rMUPs were confirmed by ESI-MS in each case. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). The figure shows the deconvoluted, true mass spectra, showing the intact masses for the rMUPs, and the true mass spectra viewed as a peak area.

Synapt G2

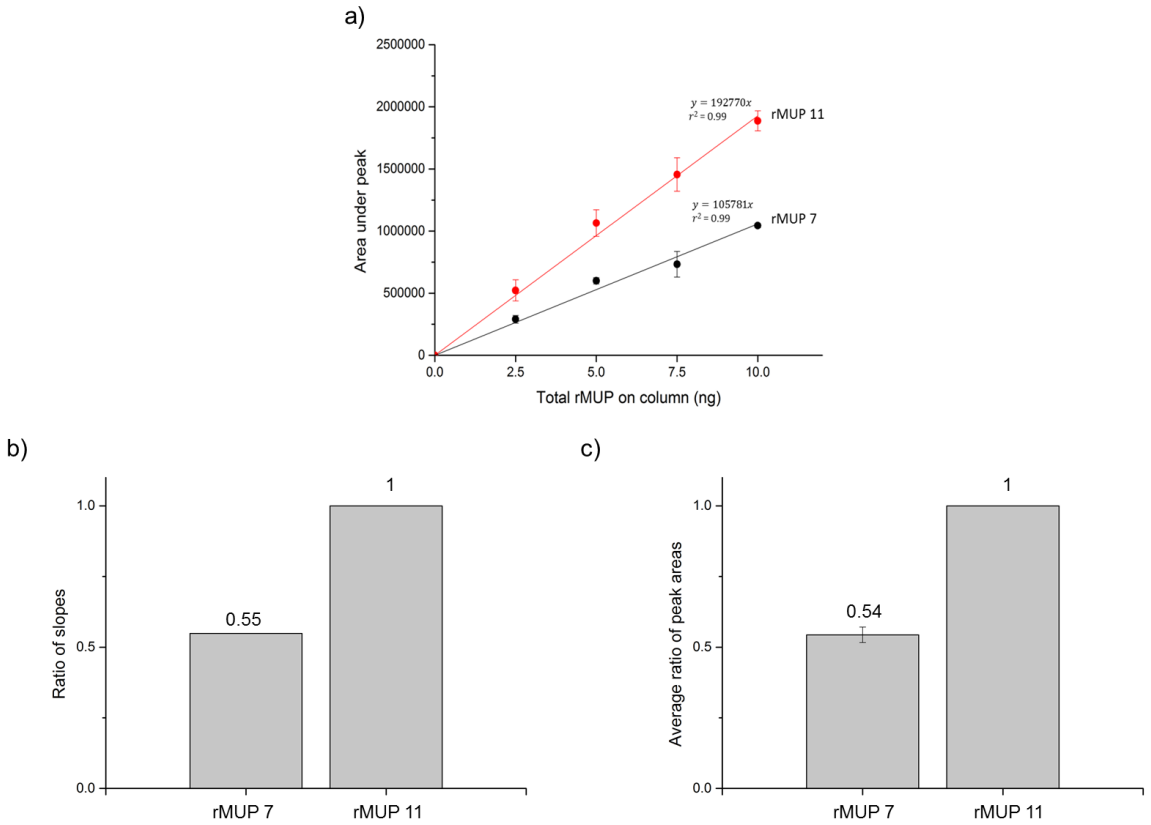


Figure 5.23 Comparison of the linearity of rMUP ESI-MS peak area with increasing protein load in a two-rMUP equimolar mixture.
a) The area under peak values for rMUPs 7 and 11 were plotted on the same scale for a comparison between the peak area/protein load relationships exhibited by each of the rMUPs in the equimolar mixture. Error bars represent SD (n = 3). b) The ratio of the slopes of the linear relationship for each rMUP, normalised to 1 (the value of the steepest gradient). c) The average ratio of peak areas for each MUP normalised to 1 (the largest MUP peak area at each protein load). n = 4 (each protein load), error ± SD.

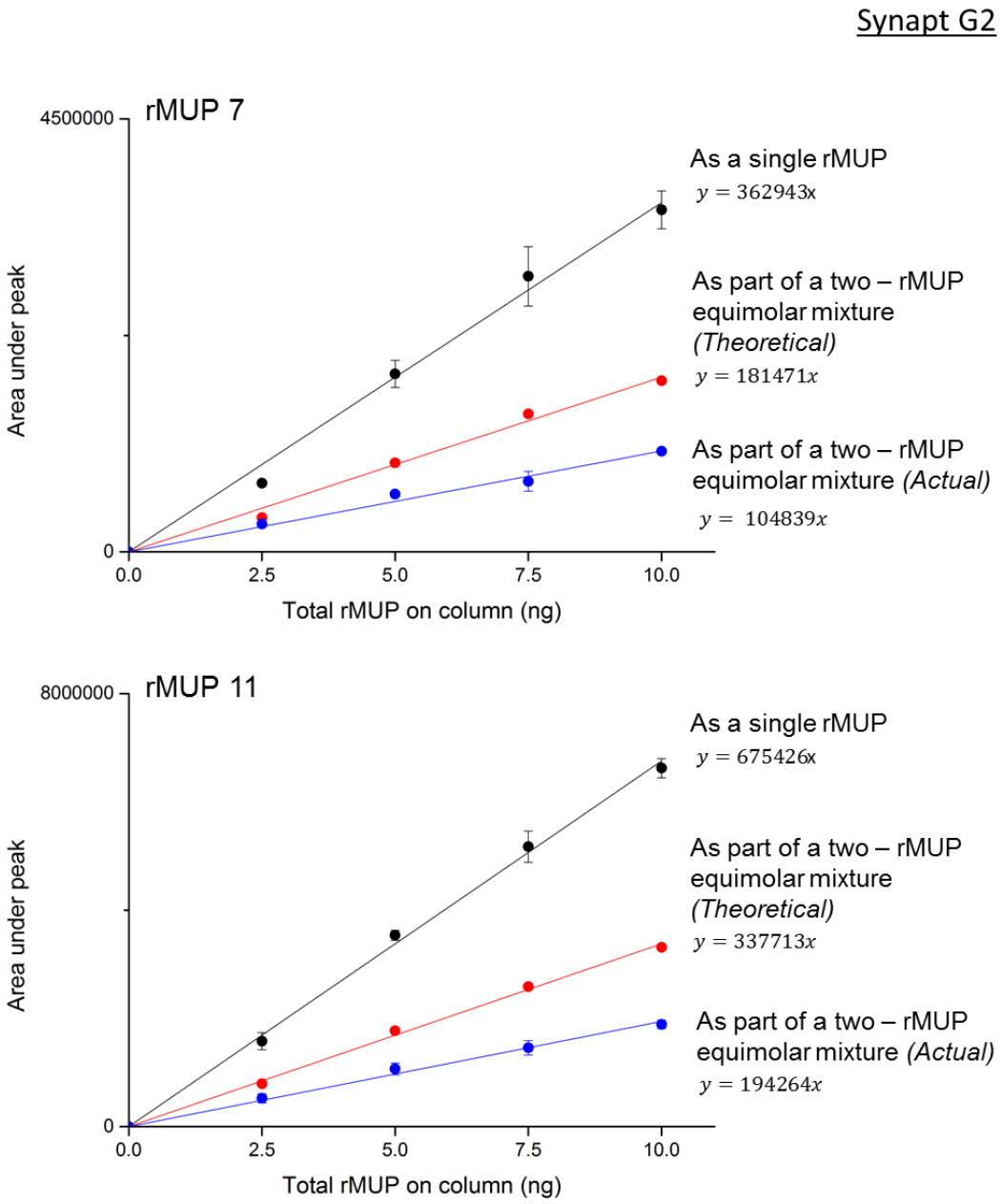


Figure 5.24 Comparing the different ESI-MS responses exhibited by each rMUP as part of a two-rMUP equimolar mixture and as a single rMUP sample.
The area under peak values for each rMUP as total protein load on column increased were recorded for the proteins as single samples (black data points) as well as part of a two-rMUP equimolar mixture (blue data points). These were plotted on the same scale, along with theoretical area under peak values expected for the rMUPs in the two-rMUP mixture if the rMUPs were to exhibit the same ESI-MS responses as in single rMUP samples (red data points). This enabled the visualisation of the differences in the theoretical and actual ESI-MS responses exhibited by each rMUP as part of an equimolar mixture. Error bars represent SD (n=3).

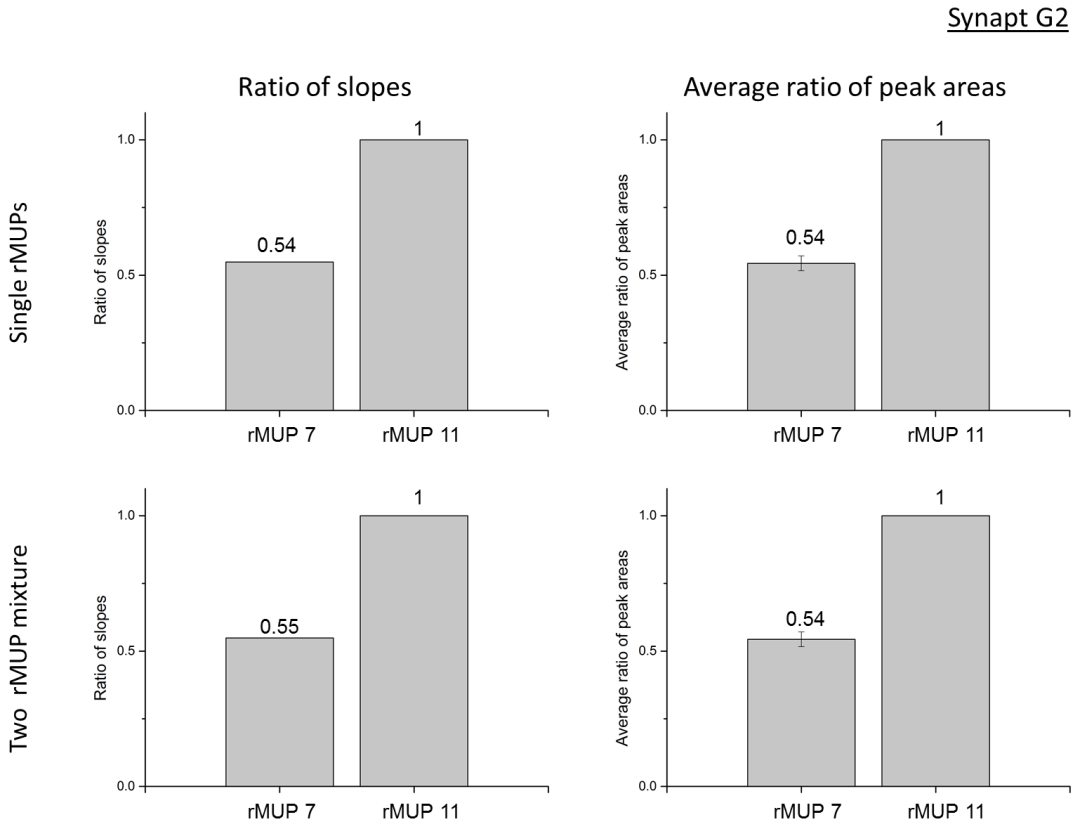


Figure 5.25 The ratio of slope and peak area values for rMUPs 7 and 11 as single samples and as part of a two-rMUP equimolar mixture.
The ratio of the slopes of the linear relationship for each rMUP, normalised to 1 (the value of the steepest gradient), were almost identical for rMUPs 7 and 11 in single samples and as part of an equimolar mixture. The average ratio of peak areas for each rMUP normalised to 1 (the largest rMUP peak area at each protein load) were identical for the rMUPs both as single samples and as part of an equimolar mixture (n = 4 (each protein load), error ± SD).

account the ratios of the slope values and the average ratio of peak areas, single-sample rMUPs 7 and 11 had a slope and peak area ratio of 0.54:1, and in the two-rMUP sample analysis, the ratio of slopes was 0.55:1 and the ratio of peak areas was 0.54:1. This further suggests rMUPs, and therefore MUPs, exhibit different ESI-MS responses dependent on sample complexity, but regardless of sample complexity, rMUPs (and therefore, also MUPs) exhibit the same relative ionisation efficiency, meaning ESI-MS analysis of intact proteins is a suitable method for the relative quantification of MUPs in urine samples (Figure 5.25).

In the analysis of the three-rMUP equimolar mixture, the observed ESI-MS response for rMUPs 7 and 11 were significantly lower than expected – at the 10 ng total protein load, the peak area value for rMUP 7 was 270 times lower than the theoretical value, and for rMUP 11, the peak area value at this load was 230 times lower than the theoretical value. Comparing this to the observed peak area values at the same total protein load for the two-rMUP mixture, where the value for rMUP 7 was only 1.7 times lower than the theoretical value, and for rMUP 11, only 1.8 times lower; there is a possibility that there is a proton deficit during the ionisation process and so proteins part of more complex mixtures have to compete more for charge and so ionise less efficiently, causing lower signal in the mass spectra.

5.2.6 ESI-MS responses of MUPs in male C57BL/6 urine using the Synapt G2 Q-ToF mass spectrometer

In the initial ESI-MS analysis of male C57BL/6 urine, which was undertaken using a Synapt G1 Q-ToF mass spectrometer, all four major MUP isoforms exhibited linear peak area/protein load ESI-MS relationships. In the analysis of an equimolar three-rMUP mixture using the Synapt G2, however, the ESI-MS responses for each rMUP in the sample were non-linear rather than linear. Whilst the lower signal observed in the two- and three-rMUP mixtures could be due to a proton deficiency during the ionisation process, a possible reason for the difference in the linearity of the responses, based on (and compared to) the results from the single-, two-and three-rMUP sample analyses, could be due to sample complexity. It would therefore be expected, due to rMUPs possessing identical biochemical properties to natural MUPs, that in a more complex sample (*i.e* urine), MUPs would also exhibit a non-linear power function relationship between peak area and protein load. To confirm that MUPs exhibit the same ESI-MS response as their recombinant counterparts in a more complex sample (containing four MUPs), and to confirm sample complexity effects the ESI-MS responses exhibited by rMUPs (and MUPs), the experiment

undertaken in Section 5.2.2 was repeated using the Waters Synapt G2 Q-ToF, and with a different urine sample.

Male C57BL/6 urine was diluted to 5 ng/µl, and a range of injection volumes (0.5 µl – 2 µl) of the sample resulted in total protein loads of 2.5, 5, 7.5 and 10 ng on column, as per the dynamic range of the instrument. Samples were injected into the system and ESI-MS analysis was carried out using the method described in Chapter 2. Each sample was analysed in triplicate. The raw mass spectra were processed as described in Chapter 2, giving the molecular weight, peak intensities and peak areas of each MUP in the sample. The mass spectra were deconvoluted using the entire multiply charged ion series to ensure a true representation of the peak area/protein load ESI-MS responses of each MUP in the urine sample (Figure 5.26). The ESI-MS intact protein peak areas of MUPs 7, 10, 11 and 20 at 2.5, 5, 7.5 and 10 ng total protein loads were plotted (Figure 5.27) and the responses of the peak area values with protein load for each of the MUPs were compared with those observed for the rMUPs in the three-rMUP equimolar mixture. As expected, analysis of the MUPs in male C57BL/6 urine using the Waters Synapt G2 confirmed that MUPs (and rMUPs) exhibit a non-linear peak area/protein load relationship in more complex samples. Each MUP in the sample exhibited a non-linear response (Figures 5.27 and 5.28). The ESI-MS response of MUP 10 relative to the other MUPs in the sample has not been confirmed by the respective rMUP in previous analyses, but due to it being a central MUP with the same number of protonatable sites as MUP 7, it is thought that it would exhibit a fairly similar response to MUPs 7 and 11. Figure 5.28 indicates that the urine sample contains more MUP 10 (18708 Da) relative to the other MUPs in the sample (y = 241627x^1.5, r^2 = 0.99). Based on the analysis of the three-rMUP equimolar mixture, the urine contains a similar amount of MUP 11 and MUP 7. Figure 5.28 suggests that the urine sample contains more MUP 11 (y = 203851x^1.3, r^2 = 0.98) than MUP 7 (y = 13279x^2.0, r^2 = 0.99), but knowledge of the differing ionisation efficiencies of these MUPs as a result of the rMUP analyses indicates that the urine contains similar amounts of these MUPs in relation to the other MUPs in the sample. Similarly, Figure 5.28 indicates that the urine contains slightly more MUP 20 (y = 51332x^1.5, r^2 = 0.99) relative to MUP 7, but again, knowledge of the ionisation efficiencies of the recombinant forms of these proteins suggests that the urine sample contains less MUP 20 relative to all other MUPs in the sample, due to rMUP 20 ionising more efficiently in ESI-MS analysis than rMUPs 7 and 11.

It has been demonstrated that MUPs, like their recombinant counterparts, exhibit a non-linear power function between peak area and protein load when being analysed as part of a complex sample using ESI-MS. The analysis of rMUPs as single samples, and as part of equimolar mixtures has confirmed that whilst ESI-MS analysis of MUPs in urine samples can provide useful information regarding the quantification of these MUPs relative to others in the sample, it is imperative that the ionisation efficiencies of each MUP in the sample is known prior to making relative quantification conclusions from the peak area data generated from ESI-MS analysis.

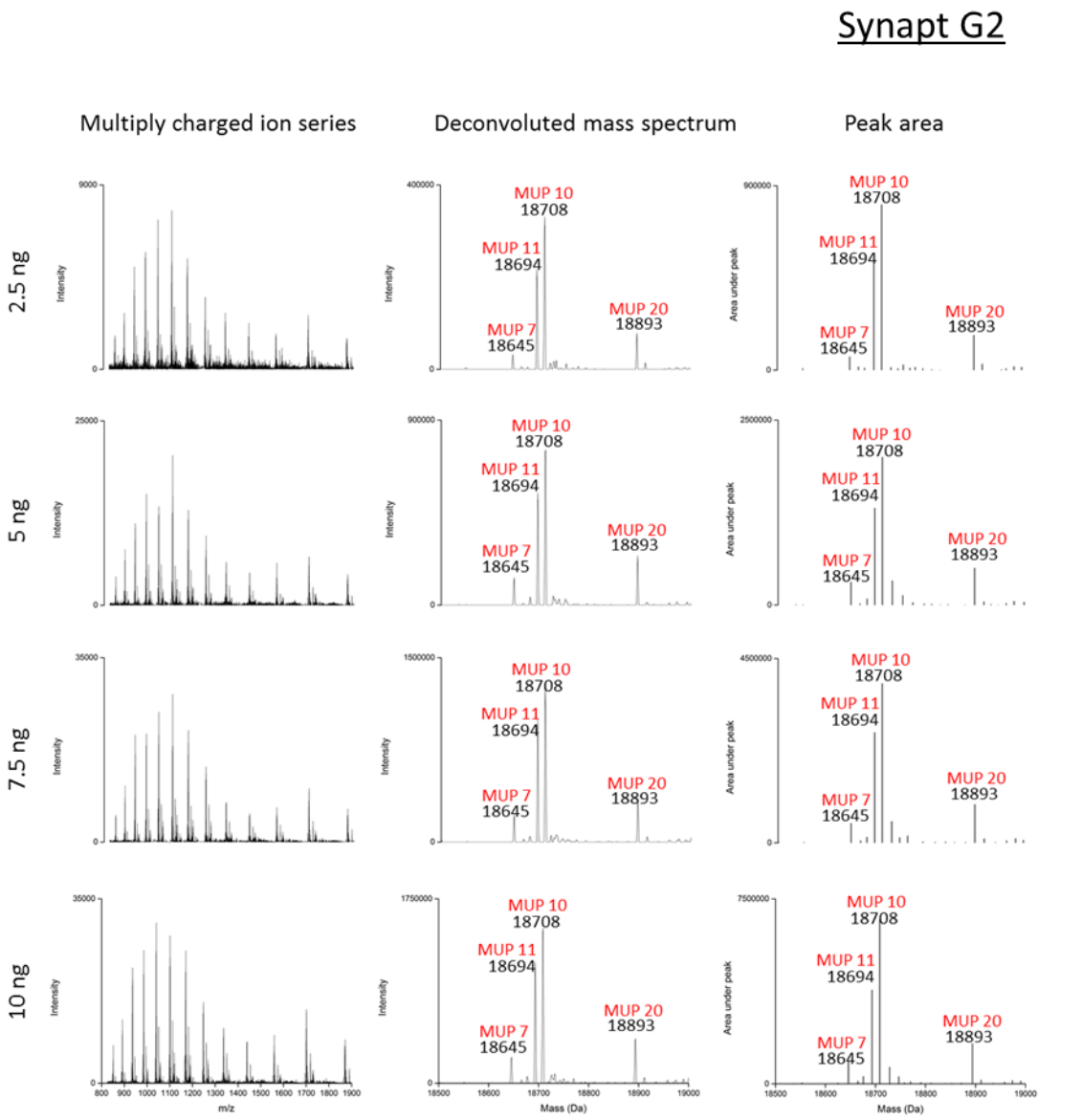


Figure 5.26 ESI-MS analysis of male C57BL/6 urine at 2.5, 5, 7.5 and 10 ng total protein loads.
Urine (diluted to 5 ng/µl in 0.1% (v/v) formic acid) was injected in triplicate over a range of injection volumes onto a C4 desalting trap and the masses of the rMUPs were confirmed by ESI-MS in each case. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). The figure shows the multiply charged ion series, the deconvoluted, true mass spectra, showing the intact masses for the rMUPs, and the true mass spectra viewed as a peak area.

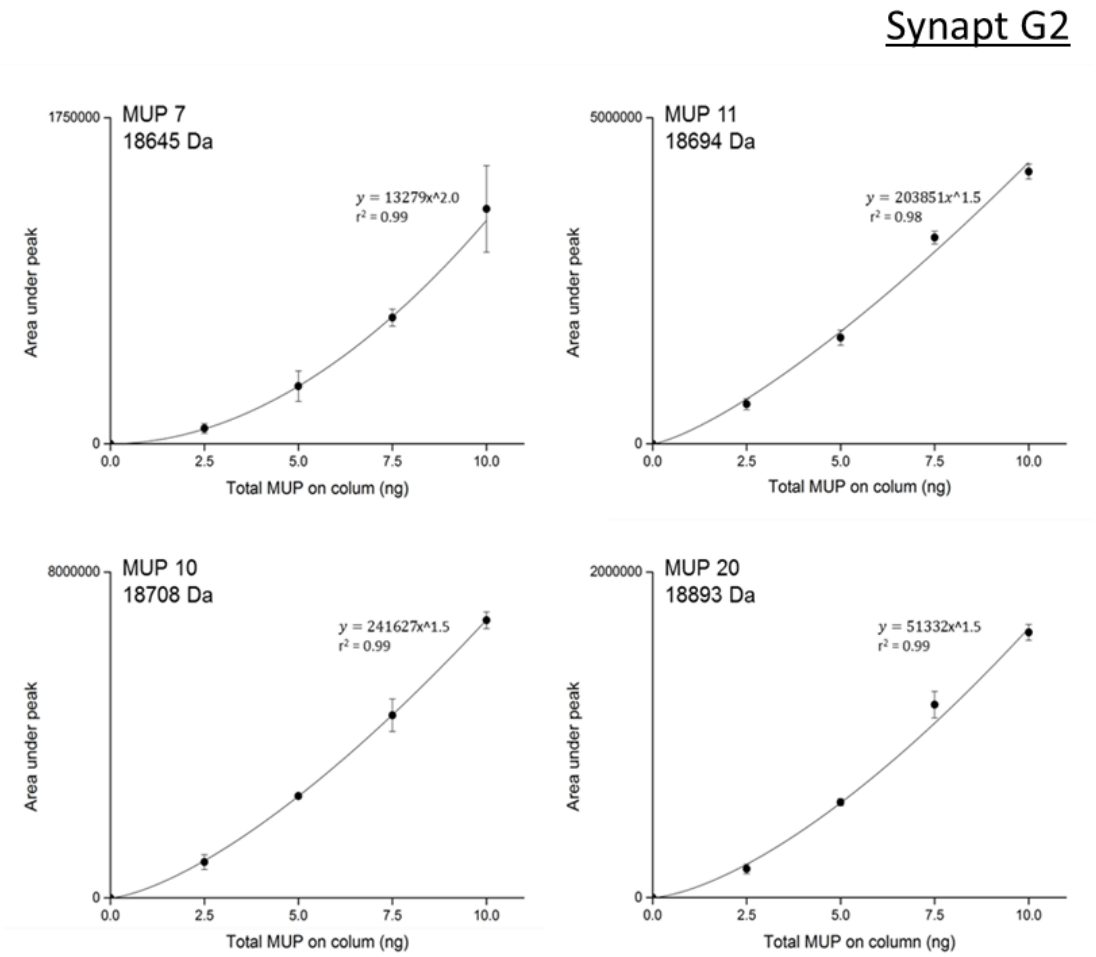


Figure 5.27 Relationship of ESI-MS peak area with increasing protein load for each MUP in male C57BL/6 urine.
Urine (diluted in 0.1% (v/v) formic acid) was injected onto a C4 desalting trap and the masses of the MUPs present were determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). Samples were injected in triplicate. The area under peak values for the four main male C57BL/6 MUP variants were assessed as total MUP load on column increased. Error bars represent SD (n = 3).

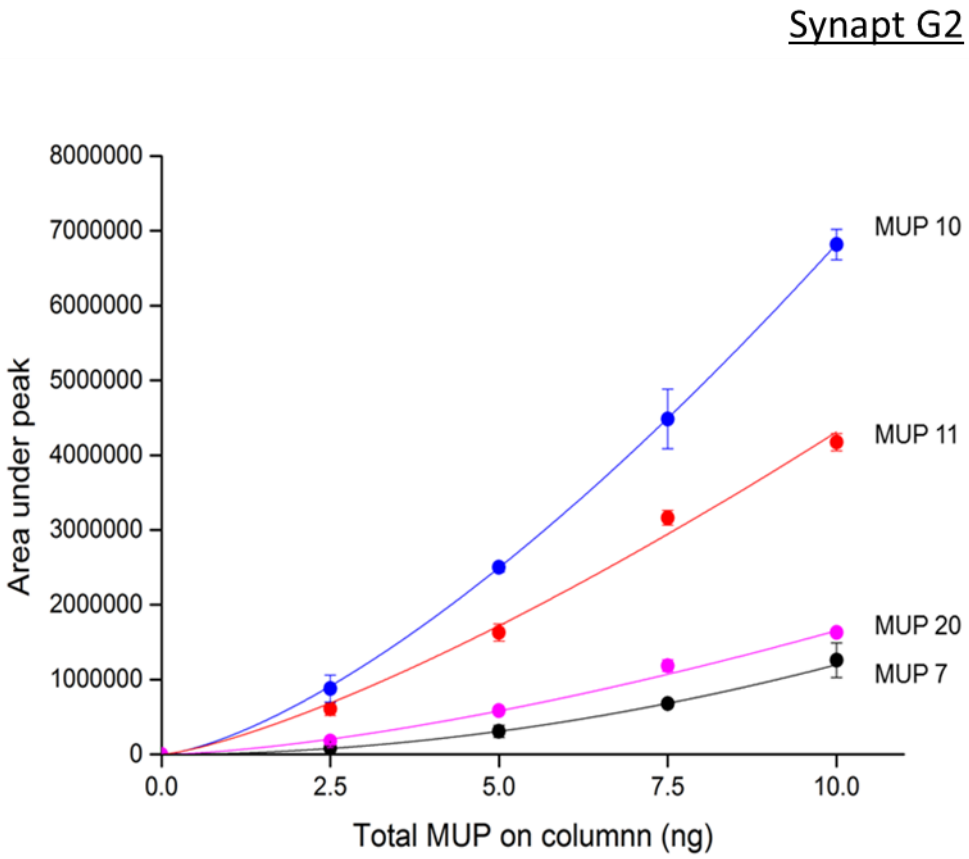


Figure 5.28 Comparison of the relationship of MUP ESI-MS peak area with increasing protein load for the quantification of MUPs relative to each other in the sample. The area under peak values for the four main male C57BL/6 MUP variants assessed were plotted on the same scale for a comparison between the peak area/protein load relationships exhibited by each of the MUPs in the urine sample. Error bars represent SD (n = 3). For relative quantification using peak areas, the ionisation efficiencies of each MUP need to be known and taken into account.

5.3 Conclusions

The aim of this Chapter was to conduct preliminary experiments to discover whether ESI-MS analysis of intact proteins, which is an established valuable method for the identification of MUPs, was a suitable method for the absolute quantification of MUPs in urine samples. Whilst MaxEnt software is able to measure molecular weight and quantitative relative intensities of the sample components in ESI-MS analysis (Ferrige *et al*. 1991,1992; Cottrell and Green 1993, 1998), knowledge of the ionisation efficiencies of each sample component is required for more accurate relative quantification data. It was determined that whilst the ionisation efficiencies of rMUPs (and therefore MUPs) can be confirmed through the ESI-MS analysis of known amounts of rMUP, with ionisation efficiencies of rMUPs remaining the same relative to others regardless of sample complexity, the relationship between peak area and protein load in ESI-MS analysis (ESI-MS response) is different for each rMUP (and therefore MUP) depending on sample complexity. There was also an indication of a proton deficiency during the ionisation process in analyses using the Synapt G2, with increasing sample complexity yielding much lower mass spectrometry signals than expected. For these reasons, it was confirmed that ESI-MS analysis of MUPs in urine would require development to be a suitable method for the absolute quantification of MUPs without prior extensive chromatographic separation of urine, but that with knowledge of each MUP's ionisation efficiency relative to others in the sample, and method development to ensure successful ionisation of all MUPs in a mixture, ESI-MS analysis of MUPs in urine is suitable for relative quantification.

Firstly, the predicted ionisation efficiencies of MUPs in ESI-MS analysis were assessed, based on differences in their number of protonatable sites, hydrophobicity and charge state distribution profiles. The similarity of MUP sequences (particularly central MUPs) and structure indicated that there were unlikely to be any major differences in ionisation efficiencies amongst them as a result of minimal differences in hydrophobicity or charge state distribution. The highly homologous central MUPs also have a very similar number of protonatable sites, again suggesting no differences in ionisation efficiencies would be observed. However, the peripheral MUPs possess more variable numbers of protonatable sites, which may result in differences in ionisation efficiencies. Evidence indicates that the number of protonatable sites a protein has, rather than its shape and molecular mass, appears to determine charge state distribution profiles for denatured proteins in ESI (Krusemark *et al.* 2009), and so this was the main consideration when assessing the relationship between MUP

concentration and observed ESI-MS response. As MUP 20 (darcin) exhibits a lower charge state distribution profile than other MUPs in ESI-MS analysis, which is consistent with it retaining a more compact structure under these experimental conditions (Phelan *et al.* 2014), the effect of charge state distribution profiles on ESI-MS response was also considered when observing ionisation efficiencies of three rMUPs – rMUPs 7, 11 and 20.

The initial experiment was to determine whether there was a linear relationship between MUP in a urine sample and the area under MUP peak in the deconvoluted, processed true mass spectrum in ESI-MS analysis of male C57BL/6 urine. A linear ESI-MS response would be required so that, in quantitative experiments with the use of rMUPs as internal standards, with knowledge of each MUP's ionisation efficiency, the measurement of the area under curve of the rMUP internal standard(s) and the corresponding MUP analyte protein(s) could provide absolutely quantitative data. This experiment was undertaken using a Waters Synapt G1 Q-ToF mass spectrometer, and it was apparent that all four major MUP isoforms retained some folding in the gas phase. The area under peaks recorded for each MUP in the male C57BL/6 urine sample increased linearly with increasing amounts of total protein on column from 0.8 ng – 3.2 ng. The relationships between protein load and area under peak of MUP 11 and MUP 20 were very similar (suggesting a similar relative amount of these proteins was present in the sample), the linear relationship for MUP 7 was shallower (indicating there was less of this protein in the sample relative to the other MUPs), and the linear relationship for MUP 10 was steeper (suggesting there was more of this MUP in the urine sample relative to all other MUPs in the urine sample). However, for absolute quantification and indeed, accurate relative quantification, knowledge of the responses of each individual MUP in ESI-MS analysis was required, which involved the assessment of the relationship of protein load and peak area whilst considering how charge state distribution, protonatable sites and sample complexity affected MUP ionisation and therefore peak area. This was explored using recombinant forms of MUPs, which allowed the analysis of individual MUPs without the need for complex and time-consuming chromatographic separation of mouse urine. It was confirmed that rMUPs 7, 11 and 20 also retained some folding under these experimental conditions, and that they too exhibited linear ESI-MS responses.

All other experiments were conducted using a Waters Synapt G2 Q-ToF mass spectrometer, as the Synapt G1 became unavailable. Once the dynamic range of this instrument was determined, recombinant MUPs 7, 11 and 20 (purified and diluted

to equal concentration, confirmed by protein assay and SDS-PAGE analysis) were diluted appropriately for ESI-MS analysis and each rMUP was analysed separately using ESI-MS to assess the relationship between rMUP protein load and peak area. Analysis of the responses of equal concentration, single-sample rMUPs in ESI-MS would determine differences in each protein's ionisation efficiency, and whether the ionisation efficiency of an rMUP relative to those of other rMUPs could be explained and confirmed by the number of protonatable sites that rMUPs possess, and/or the charge state distribution profile exhibited in ESI-MS analysis. rMUP 20 exhibited a lower degree of charging than rMUPs 7 and 11 due to its more compact structure, and rMUP 7 possessed one more protonatable site than rMUPs 11 and 20, so if these factors were to affect ionisation efficiency, it would be expected that rMUP 7 would have a slightly higher ionisation efficiency than rMUP 11, and the ionisation of rMUP 20 would be lower than both rMUPs 7 and 11. However, the results proved the opposite – rMUP 20 had a slightly higher ionisation efficiency than rMUP 11, and the ionisation efficiency of rMUP 7 was lower than both. This confirmed that whilst MUPs do have differing ionisation efficiencies in ESI-MS analysis, the efficiencies cannot be explained/confirmed by assessing protonatable sites or charge state distributions.

For ESI-MS analysis to be suitable for the relative quantification of MUPs in urine samples, the ionisation efficiencies of rMUPs relative to one another would need to be the same in samples of greater complexity, and for the absolute quantification, rMUPs (and therefore MUPs) would need to exhibit the same peak area/protein load ESI-MS response in samples of greater complexity. It was confirmed through the ESI-MS analysis of two-rMUP and three-rMUP equimolar samples that whilst the rMUPs ionisation efficiencies relative to one another remained unchanged in samples of differing complexities, the peak area/protein load responses exhibited by all three rMUPs changed depending on sample complexity. It was therefore concluded that whilst ESI-MS analysis of MUPs in urine could be suitable for relative quantification, given that the ionisation efficiencies of all MUPs in the sample relative to one another were known, the differing relationships between peak area and protein load in ESI-MS analysis (ESI-MS response) for each rMUP in different sample complexities meant rMUP spiked into urine in known amounts could not produce suitable standard curves for absolute quantification. The only way that ESI-MS analysis of mouse urine with rMUPs as internal standards could be suitable for absolute quantification would be if a time consuming method chromatographic method was developed that was able to separate each MUP and its spiked rMUP counterpart into a single fraction prior to ESI-MS analysis. The chromtagraphic separation of urine prior to spiking the

required rMUP into each MUP fraction would be a simpler method to develop, but unsuitable since losses of MUPs in the sample in all prior preparation steps would not be accounted for.

The ESI-MS analysis of male C57BL/6 urine, with prior knowledge of the ionisation efficiencies of three of the four major MUP isoforms, confirmed the importance of the knowledge of the ionisation efficiencies of each MUP in the sample for accurate relative quantification. In the urine sample analysed, if relative quantification was to take place without prior knowledge of each MUP's ionisation efficiencies relative to one another, MUP peak areas (as determined by the MaxENT software) would determine the urine content as MUP 10 > MUP 11 > MUP 20 > MUP 7, but knowledge of the MUP 7, 11 and 20 ionisation efficiencies indicates that MUP 10 > MUP 11 = MUP 7 > MUP 20 is more accurate relative quantification.

For more accurate relative quantification of the MUPs in male C57BL/6 urine, the ionisation efficiency of MUP 10 in relation to the other MUPs in the sample would need to be determined. This would be done by assessing the peak area/protein load ESI-MS response using a recombinant form of this protein in the same way as for rMUPs 7, 11 and 20. Knowledge of the ionisation efficiencies of these four MUPs would allow the accurate relative quantification of MUPs in the urine of male and female C57BL/6 and BALB/c (Mudge *et al.* 2008). Further experiments would involve assessing the instrument responses observed in the Synapt G2 –is the low response for rMUP mixtures due to a proton deficiency in the ionisation process? Do recombinant MUPs with the hexa-histidine tag removed (and therefore identical to natural MUPs) exhibit this same low response in an equimolar mixture? The use of a stronger acid in sample preparation would address whether this issue was as a result of proton deficiency, and the endoprotease trypsin can be used to cleave the hexa-histidine tag from purified rMUPs, allowing a comparison of the ESI-MS responses of rMUPs with and without the tag. By conducting various experiments to ascertain why ESI-MS responses are a) lower than expected and b) non-linear should enable the identification of the causes and therefore allow the current ESI-MS method to be developed to ensure linear ESI-MS responses are exhibited by MUPs, for more accurate relative quantification of MUPs in samples of differing complexities. This in turn may allow the identification of more significant differences in relative MUP expression patterns between an individual's urine samples that may arise due to differences in situations such as reproductive status and social setting. In wild mice, MUP profiles are used by conspecifics in recognition and to avoid inbreeding. In the

urine samples of wild mice, the MUPs expressed would need to be identified prior to relative quantification so the ionisation efficiencies of all the MUPs could be confirmed to allow accurate relative quantification. It is important to remember, however, that current ESI-MS analysis of intact MUPs is unable to identify some MUP isoforms as the sequence homology of MUPs means the mass spectrometer is unable to resolve MUPs of similar mass (for example, the 18693 and 18694 Da MUPs), suggesting that for absolute quantification of MUPs in urine, peptide-level analysis is required.

For mass spectrometry based quantification of proteins larger than around 15 kDa, such as MUPs, a 'bottom up' workflow is usually required, as the generated peptides are limited in the number of protonatable sites and usually have an *m/z* value that is within the detection range of most mass spectrometers. As discussed previously, a MUP QconCAT was developed for the absolute quantification of MUPs, but problems including the high sequence similarity between the MUP variants, and poorly ionising peptides resulted in a limited choice of peptides suitable for absolute quantification. Although the implementation of a subtraction method allowed the quantification of MUPs in male and female C57BL/6 urine, the use of a multiple reaction monitoring (MRM) method with a triple quadrupole mass spectrometer could improve the quantification of MUPs in urine using QconCAT methodology, providing increased sensitivity and a greater dynamic range for the detection and quantification of less abundant MUP variants in complex sample mixtures (Lange *et al*. 2008). Despite this, the use of the MUP QconCAT faces further limitations arising from the high evolution rate of MUPs. It is possible that new MUP isoforms or allelic variants could contain peptides not represented by the current MUP QconCAT, meaning that further QconCATs would need to be designed (Beynon *et al*. 2014).

The relative and absolute quantification of MUP isoforms in the urine samples of inbred laboratory strains of the house mouse, and some isoforms expressed in the urine of wild-caught mice, can be achieved by the ESI-MS analysis of intact proteins and by the use of a MUP QconCAT. For further MUP isoforms expressed by wild-caught mice, however, quantification methods would need to be developed further, either by confirming the ionisation efficiencies of these proteins (for relative quantification) or through the development of a new QconCAT based strategy (for absolute quantification).

Chapter 6: Conclusions

6.1 General conclusions

The overall aim of this thesis was to apply proteomic techniques, primarily mass spectrometry, to the understanding of behaviour in *Mus* species. The analysis of major urinary proteins (MUPs), whose function is chemical signalling amongst various *Mus* species, enabled the discovery of MUPs in the urine of a *Mus* species where no genome sequence was available. The analysis of MUPs also allowed the tracking of investment in communally nursing BALB/c mice, and ESI-MS analysis of *M. m. domesticus* MUPs determined whether intact protein mass spectrometry, established for the identification of MUP isoforms, was suitable for the quantification of MUPs.

M. spicilegus are genetically close to the *M. musculus* group of the *Mus* species group, yet display very different mating and social behaviours to those observed in the house mouse, *M. m. domesticus*. *M. spicilegus* female mice show a preference for one particular mate, which is suggestive of a monogamous mating system, whilst the house mouse exhibits polygamous mating behaviour (Patris and Baudoin 1998; Patris *et al*. 2000). Female house mice have a sexual preference for dominant males and have developed a communal nesting system, whereas *M. spicilegus* are not cooperative breeders; they exhibit strong mating pair bonds and females appear aggressive towards one another (Tong *et al*. 2012). Since urinary MUPs play a significant role in the reproductive and social behaviours displayed by the house mouse, Chapter 3 of this thesis evaluated *M. spicilegus* urinary MUP content and explored possible roles of these proteins in scent communication. Four MUPs were identified and sequenced in the urine of *M. spicilegus* mice in the absence of genomic data, three of which were male-specific. SDS-PAGE analysis, protein assays and creatinine assays confirmed male mice expressed more MUP in their urine than females, and that the amount of MUP expressed by different males varied quite significantly whilst female MUP expression was almost identical amongst individuals. ESI-MS profiling of the urine of a number of males and females highlighted two dominant MUP masses present in all male samples – 18742 Da and 18762 Da. In some male urine samples, masses of 18585 Da and 18918 Da were identified, however all female samples only contained the 18918 Da mass, with no evidence of the 18742 Da, 18762 Da and 18585 Da masses. This indicated that these three MUPs were male-specific, whilst the 18918 Da MUP was not sex specific. The

molecular weights of the four identified *M. spicilegus* MUPs did not match the molecular weights of any previously identified house mouse MUPs, therefore further mass spectrometric analysis was required for the identification of the amino acid sequences of the *M. spicilegus* MUPs.

To determine full amino acid sequences for each of the *M. spicilegus* MUPs, the male-specific MUPs were separated using strong anion exchange chromatography prior to digestion and *de novo* sequencing analysis via LC-MS/MS and PEAKS software. Full peptide sequence information for the three male-specific MUPs was acquired, however, only the majority of the non-sex specific 18918 Da MUP could be sequenced. All *M. spicilegus* MUPs were aligned with all 21 *M. m. domesticus* MUPs and phylogenetic analysis was performed. The three male specific *M. spicilegus* MUPs were very similar to central *M. m. domesticus* MUPs, with 18742 Da and 18585 Da MUPs being very similar to the male-specific *M. m. domesticus* MUP 7 in particular. The 18762 Da MUP is most similar to *M. m domesticus* MUPs 13 and 17, and the non-sex specific 18918 Da MUP is very similar to the peripheral *M. m. domesticus* MUP 6.

The presence of male specific MUPs in male urine, along with the variance in MUP expression seen in male urine, suggested that these MUPs may have a functional role in the unusual sexual and social behaviour observed in *M. spicilegus*. Experiments were set up to determine whether contact between males and females, breeding between males and females and contact between males caused differences in their MUP expression output in the aim of determining a possible role of MUPs in their behaviour. Female mice did not alter their MUP expression in any situation, but male mice increased their MUP expression fairly significantly upon contact with females. Of particular interest was the increase in male specific 18762 Da MUP expression in male urine upon contact with females, but after male and female pairs were put together for breeding, all five males reduced their MUP expression, three of them significantly, including that of the 18762 Da MUP. This indicated that this MUP may play a role in the attraction of females. In male/male pairs, the only pair of mice that appeared to significantly alter their MUP output during contact and after interactions were the mice that displayed significant aggressive behaviour during the interactions. One male in this pair, after interactions, significantly increased their MUP output; the other male significantly reduced their MUP output after interactions, along with their expression of the male-specific 18742 Da and 18762 Da MUPs,

suggesting that increased MUP output (especially of these two male-specific MUPs) may also have a role in signalling dominance.

Whilst the determination of the roles of MUPs in scent communication in *M. spicilegus* is in its early stages, the role of MUPs in scent communication including modulating identity signalling, attraction and aggressive behaviour in *M. m. domesticus* is well established. It has been confirmed very recently (Green *et al*. 2015, published October 2015) that MUPs have a role in kin recognition in the context of communal nursing in female house mice, who prefer to nest with partners sharing their own MUP genotype. Previous studies had suggested that kin recognition and selection were important factors in the communal nesting and nursing behaviour displayed be female house mice – it was confirmed that communal nursing increases reproductive success, regardless of the relatedness of female pairs (Konig 2006). However, mutualistic cooperation has been seen to be higher in nests where females are related, meaning females may invest more in the offspring of their closer relatives (*i.e*. sisters) than their more distant relatives (cousins). It may be that females do not discriminate between their offspring and their non-offspring, and invest cooperatively more in larger, healthier pups. In Chapter 4, whole animal metabolic labelling with stable isotope labelled amino acids was used to determine the proportion of lactation investment from communally nursing female house mice and to ascertain any discriminatory factors, such as the relatedness of the communally nursing females.

In complex systems such as animals, protein turnover studies using stable isotope labelling involves the introduction of labelled amino acids via diet, however, incorporation of the labelled amino acid cannot be instantaneous due to the biomass of the animal tissues contributing to the precursor pool. Claydon *et al.* (2012) designed an experiment to incorporate labelled amino acids into mice via a 50% labelled (therefore palatable to the mice) rodent diet, using MUPs in urine to monitor the labelling of the liver - there is no evidence of MUP uptake during renal filtration, and so MUPs reflect the properties of the precursor RIA in the liver. This experimental technique was deemed suitable to be used in determining the proportion of lactation investment from communally nursing female mice – female mice were fed labelled diets, and the incorporation of labelling was determined by mass spectrometric analysis of MUPs in urine. The resultant labelled/unlabelled MUP peptides behaved the same under mass spectrometric analysis whilst providing a mass difference to differentiate between isotopes for the calculation of label incorporation. Label incorporation was calculated from precursor relative isotopic abundances (RIAs):

$$RIA = I_H / (I_H + I_L)$$

where I_H is the signal intensity displayed by the heavy (labelled) peptide and I_L is the signal intensity of the light (unlabelled) peptide.

In the communal nursing experiment, female mice, feeding their pups, passed their labelled proteins onto their pups, which were then incorporated and tracked in their offspring. LC-MS analysis of digested proteins from homogenised pup tissue samples were used to determine the proportion of lactation investment from mothers; label incorporation, and therefore investment, was observed in litters of pups where their mothers were related and in litters where their mothers were unrelated, to determine any discrimination in investment in relation to the relatedness of the female pairs in the nest. The initial preparatory experiment confirmed the rate of incorporation of labelled amino acids into adult female BALB/c mice via a 50% labelled diet by analysing MUP peptides from daily urine samples. The testing of a feeding mechanism followed, that would enable two communally nesting females to exclusively consume differently labelled diets, allowing their investment in the pups to be tracked via the transfer of two different labels from the mothers to the pups. An initial communal nursing experiment, with three pairs of females consuming the same labelled diet, confirmed that the labelled amino acids were being passed from mothers to pups over the course of six days. MALDI-ToF-MS and LC-MS analysis of digested milk proteins from the pups' stomach contents confirmed labelled amino acids were being passed from mothers to pups, and LC-MS analysis of digested proteins from pup liver and muscle samples confirmed that the labelled amino acids were being successfully incorporated into pup tissues. The aim was to find a high and low turnover protein from each tissue type, with high turnover proteins incorporating the label more quickly (high precursor RIAs) than low turnover proteins (low precursor RIA); proteins with different turnover rates in pup tissues were hoped to give information regarding investment from mothers at different time points during the final communal nursing experiment. High turnover proteins degrade and return to the protein precursor pool quickly, with newly synthesised labelled proteins taking their place. This high turnover of proteins cause rapid labelling of the precursor pool, so recent changes in label investment could be identified through the analysis of these proteins. The opposite is true for low turnover proteins, meaning that these proteins may be able to report on variations in label investment over a longer period of time. The precursor RIAs of proteins with apparent high and low turnover rates were observed in the tissues over the course of the initial communal nursing experiment to

determine whether these proteins were indeed high/low turnover and whether they could be used to report variations in labelling at different points during the later communal nursing. Whilst these different proteins incorporated the label at different rates and reached different maximum RIAs by the end of the experiment, indicating that these proteins had different turnover rates, none of the proteins reached a maximum labelling plateau. The slow incorporation of label into pup proteins is due their rate of growth; however, all four of these proteins in pup tissues labelled at a significant enough rate to be suitable for analysis of pup tissues to track investment from differently labelled females in the final communal nursing experiment.

For the communal nursing study, 8 related female pairs and 9 unrelated female pairs were set up to give birth and nurse their litters in the communal nest. In order to track the investment from each mother in a communal litter, each mother consumed a differently labelled diet (one labelled with [D_4] lysine, the other with [D_9] lysine) via the established feeding mechanism when the first litter in the nest to be born was 7 days old. The mothers were fed these diets for 7 days, resulting in two different labels being incorporated, one into each mother, which were passed on to the communal litter via their milk, and incorporated into the pups they invested in. The precursor RIAs of each label present in the proteins in the pup's tissues indicated the proportion of investment from each mother. Of the 8 related pairs, 7 had communal litters; of the unrelated pairs, 6 had communal litters. In the remaining four pairs, only one female had a litter that survived to the labelling stage. Firstly, the MUPs in the mothers' urine sampled on the final day of the experiment were analysed using LC-MS to confirm the incorporation of label via diet. In the 7 related pairs, only the females assigned the [D_4] lysine labelled diet were fully labelled with the correct diet, with no evidence of consuming any of the unassigned [D_9] lysine diet. Only three out of seven females assigned the [D_9] lysine diet were fully labelled with the correct diet; four were partially [D_9] lysine labelled, and two of these showed evidence of having consumed some of the unassigned [D_4] lysine diet. The reason for these mice having lower [D_9] lysine RIA values was likely to be as a result of them consuming less diet; it is unknown whether these mice were consuming less diet because of a possible problem with the feeding mechanism, however this is unlikely due to no problems being encountered in any of the previous experiments. The fact that two of the [D_9] lysine assigned mice had very low intensity [D_4] lysine peptides in their urine could be due to the [D_4] lysine female bringing and dropping small amounts of their [D_4] lysine diet into the communal nest, and so accessible by the [D_9] lysine assigned female. In the 6 unrelated pairs, all females were fully labelled with the correct diet, with no

evidence of having consumed any unassigned diet, further suggesting that in the related pairs, the feeding mechanism was functioning correctly. LC-MS analysis of the digested pup liver and muscle proteins enabled the precursor RIA values for [D_4] and [D_9] lysine labelled peptides to be calculated for the four proteins (selected in the milk labelling study) in the tissue samples of pups from litters where both mothers were fully, correctly labelled. The difference between group mean RIAs and the variation among and between group RIAs was assessed for each of the four proteins using analysis of variance (ANOVA) or in cases where litter sizes were too small for ANOVA, the Welch two sample t-test. To determine whether either female invested significantly more (or less) in the entire communal litter, the difference between entire group mean RIAs was assessed using the Welch two sample t-test. This determined whether there was any statistical significance between the investment from a female in her own pups and in her partner's pups. There were no statistically significant differences between a mother's investment in her own pups and her partners in all communal litters where mothers were related (*i.e*. no difference between the [D_4] lysine RIAs calculated in the pups belonging to the [D_4] lysine mother and the pups belonging to the [D_9] lysine mother, and *vice versa*); there were also no significant differences between a mother's investment in her own pups and her partners in communal litters where mothers were unrelated. However, in a number of communal litters, both with related and unrelated female pairs, one female appeared to invest more in the entire communal litter than the other. The reason for this cannot be determined due to the small number of communal litters studied. There was no definitive link between size or age of litter and the proportion of communal investment from their mother, but when relating back to previous literature, Wilkinson and Baker, 1988 and Konig 1989 state that in captivity, wild female house mice appear to take turns nursing each other's offspring in communal nests. Taking into account the growth rate and high protein turnover in pups, it is possible that in this communal nursing study, the females have been taking turns in indiscriminately investing in the entire communal litter, and it could be that the female who appeared to invest 'more' was in fact the last female to nurse the communal litter before the end of the experiment, and that throughout the experiment, both females were investing similarly, taking it in turns to nurse the entire communal litter.

The lactation investment in the four non-communal litters was also studied; in three pairs, only the mother of the pups in each pair was fully labelled. In one pair, both females were fully labelled, and in all four pairs, no female appeared to consume any of the unassigned diet. The fact that the females with no surviving pups generally

appeared partially labelled may be due to them consuming less diet due to stress, or the fact they did not need to consume as much since they were not expending energy investing in the pups. In all four litters, pups received investment from their mothers only, and none from the other female in the communal nest. This was the case in both the related pair and the unrelated pairs. As a result, the differences in investment from each mother in the litter was highly significant.

In this Chapter, the proportion of lactation investment from communally nursing house mice in pup litters was successfully studied with the use of two different labels, and it was determined that both related and unrelated female pairs do not discriminate between nursing their own litters and the litter from the other female. However, in pairs where only one female's pups survived, only the mother of the litter invested in the pups, with the other lactating female not investing in any of the other female's pups. The fact that in most communal nests one mother invested in the entire communal litter than the other meant that it would be of interest to determine whether litter age differences linked to the a female investing more than the other in the communal litter, or whether the apparent differences in investment are actually due to the mothers taking turns in investing in the entire communal litter.

The analysis of MUPs, particularly using mass spectrometry, has been used for determining labelled amino acid incorporation into inbred female house mice as well as for testing the feeding mechanism for the communal nursing experiment. The analysis of mothers' urinary MUPs in the communal nursing experiments was instrumental for the definitive assessment of the proportion of labelled amino acid, and therefore investment, pups in a communal litter received from each mother. Mass spectrometry enabled the identification and characterisation of new MUPs in *M. spicilegus*, with the analysis of MUP expression patterns using ESI-MS in different social situations providing an insight into their possible roles in scent communication. Whilst ESI-MS of intact proteins has proved invaluable for the identification of the molecular weights of most MUP isoforms, it is essential that as well as identifying the MUPs and their roles in chemical signalling, the differences in MUP expression profiles are quantified. For a definitive assessment of the variation in MUP expression of an individual in different situations, and the differences in MUP expression profiles between individuals, a suitable quantification method is required.

Previously, the absolute quantification of MUPs was based on QconCAT technology, however, problems were encountered relating to the high sequence homology of MUPs. Though the MUP QconCAT was effective for the quantification of the MUPs

expressed by an inbred laboratory strain, further complications could arise from the high rate of evolution of these proteins; new allelic variants of MUPs may generate peptides different to those in the designed MUP QconCAT upon proteolysis. The analysis of MUPs from *M. spicilegus* has revealed MUPs that are different in sequence to those present in the C57BL/6 laboratory mouse strain, and in both of these cases, new QconCATs would have to be designed and constructed. Since the current challenges with MUP quantification are mainly based on the proteins' sequence homology, the high rate of MUP evolution and the lack of a unique peptide for each protein, the use of a full length protein standard may address these problems. It was considered whether quantification could take place at the intact protein level ('top down'), since intact protein analysis by ESI-MS is already well established for the accurate identification of MUP isoforms. ESI-MS can also provide an indication of the amount of each MUP present in the sample relative to each other, however, it cannot be used to give an accurate representation of the amount of each MUP isoform present in a sample. This is because although MUPs are very similar in sequence, they may ionise differently to each other in the source during ESI-MS that may affect the accuracy in the resulting relative quantitation data. For this reason, adding a standard that has identical biochemical properties to its target but with the addition of a label in order to distinguish it from the analyte in mass spectrometry analysis, in a known amount to a sample prior to intact mass analysis by ESI-MS would be desirable, in conjunction with a standard curve.

Chapter 5 outlined preliminary experiments to investigate whether ESI-MS analysis of MUPs, in conjunction with the use of known amounts of rMUPs as standards, can allow absolute quantification of MUPs. The ionisation efficiencies of MUPs relative to one another in samples of differing complexity were also assessed to confirm that ESI-MS analysis of intact proteins was also suitable for the relative quantification of MUPs in urine samples. Firstly, it was determined that small differences in the number of protonatable and hydrophobic sites MUPs possessed (due to their sequence homology) were unlikely to result in differences in ionisation efficiency. The same was true for charge state distributions, with the exception of MUP 20 (darcin), which retains a more folded structure in the gas phase and so exhibits a lower charge state distribution profile, suggesting this MUP may generate a lower intensity mass spectrometry signal. The linearity of MUP ESI-MS response (the relationship between protein load and area under MUP peak in the processed true mass spectrum) was revisited and confirmed using the Synapt G1 mass spectrometer, since a linear response would be required if ESI-MS was to be used as a quantitative

method. This was explored using male C57BL/6 urine and rMUPs, which needed to exhibit identical ESI-MS behaviour to their native counterparts if they were to be suitable absolute quantification standards, and this was the case, with both natural and recombinant MUPs exhibiting linear ESI-MS responses and two charge state distribution profiles, indicative of all MUPs retaining a second, more folded conformational state under these experimental conditions.

All other preliminary experiments took place using a Synapt G2 mass spectrometer, since the G1 become unavailable. To determine any differences in ionisation efficiency between MUPs, and whether they could be explained by charge state distribution profiles, rMUP samples (7, 11 and 20) were diluted to equal concentration (confirmed by SDS-PAGE and protein assay) and separately analysed using ESI-MS to assess the relationship between protein load and peak area. In Synapt G1 analysis, rMUP 20 exhibited a lower degree of charging than rMUPs 7 and 11 due to its more compact structure, it would be expected that the ionisation efficiency of rMUP 20 would be lower than both rMUPs 7 and 11. However, the results proved the opposite – rMUP 20 had a slightly higher ionisation efficiency than rMUP 11, and the ionisation efficiency of rMUP 7 was lower than both. Whilst rMUPs 7 and 11 exhibited linear ESI-MS responses as expected, rMUP 20 exhibited a non-linear response where peak area increased disproportionately with peak load. This experiment confirmed that whilst MUPs do have differing ionisation efficiencies in ESI-MS analysis, the efficiencies cannot be explained/confirmed by assessing charge state distributions, and that rMUP 20 exhibited a different response in Synapt G2 analysis than in Synapt G1 analysis.

For ESI-MS analysis to be suitable for the relative quantification of MUPs in urine samples, the ionisation efficiencies of rMUPs relative to one another were required to be the same in samples of greater complexity, and for the absolute quantification, rMUPs (and therefore MUPs) would need to exhibit the same peak area/protein load ESI-MS response in samples of greater complexity. It was confirmed through the ESI-MS analysis of two-rMUP and three-rMUP equimolar samples that whilst the rMUPs ionisation efficiencies relative to one another remained unchanged in samples of differing complexities, the peak area/protein load responses exhibited by all three rMUPs changed depending on sample complexity. In three-rMUP samples, rMUPs 7 and 11 exhibited non-linear ESI-MS responses, when as single samples and as part of two-rMUP samples, they exhibited linear responses. In the analysis of the three-rMUP equimolar mixture, the observed ESI-MS response for rMUPs 7 and 11 were

significantly lower than expected – at the 10 ng total protein load, the peak area value for rMUP 7 was 270 times lower than the theoretical value, and for rMUP 11, the peak area value at this load was 230 times lower than the theoretical value. Comparing this to the observed peak area values at the same total protein load for the two-rMUP mixture, where the value for rMUP 7 was only 1.7 times lower than the theoretical value, and for rMUP 11, only 1.8 times lower; there is a possibility that there was a proton deficit during the ionisation process and so proteins part of more complex mixtures have to compete more for charge and so ionise less efficiently, causing lower signal in the mass spectra.

The ESI-MS analysis of male C57BL/6 urine, with prior knowledge of the ionisation efficiencies of three of the four major MUP isoforms, confirmed the importance of the knowledge of the ionisation efficiencies of each MUP in the sample for accurate relative quantification. In the urine sample analysed, if relative quantification was to take place without prior knowledge of each MUP's ionisation efficiencies relative to one another, MUP peak areas would determine the urine content as MUP 10 > MUP 11 > MUP 20 > MUP 7, but knowledge of the MUP 7, 11 and 20 ionisation efficiencies indicates that MUP 10 > MUP 11 = MUP 7 > MUP 20 is more accurate relative quantification.

The preliminary work outlined in Chapter 5 highlights the unpredictable behaviour of rMUPs (and therefore MUPs) in the gas phase, as well as the fact that although highly homologous, different rMUPs have different ionisation efficiencies in ESI-MS analysis that cannot be explained by charge state distribution profiles. The fact that rMUPs have the same ionisation efficiencies relative to one another in samples of differing complexity, however, indicates that ESI-MS analysis is suitable for the relative quantification of MUPs in mouse urine, given that the ionisation efficiencies of all MUPs in the sample relative to one another are known and considered in data analysis. Developing the ESI-MS method could increase the accuracy of the relative quantification data and may allow the development of a protein-level absolute quantification method that does not require prior extensive chromatographic separation of urine.

6.2 Future work

For more accurate relative quantification of the MUPs in male C57BL/6 urine, knowledge of the ionisation efficiency of MUP 10 in relation to the other MUPs would be required, by assessing the ESI-MS response of this protein using the respective rMUP. Knowledge of the ionisation efficiencies of MUPs 7, 10, 11 and 20 would allow the accurate relative quantification of MUPs in the urine of male and female C57BL/6 and BALB/c . Further experiments would involve assessing the instrument responses observed in the Synapt G2 – to determine whether low rMUP responses in mixtures was due to a proton deficiency in the ionisation process, the use of a stronger acid in sample preparation would confirm whether this was the case. A comparison of the ESI-MS responses of rMUPs with the hexa-histidine purification tag removed (using trypsin) with the rMUPs analysed in Chapter 5 would determine whether the tagged rMUP exhibited identical ESI-MS behaviour and responses to the respective untagged rMUP (which is identical to the native MUP), to confirm whether an absolute quantification method could be developed using tagged rMUPs as a standard. Conducting various experiments to ascertain why ESI-MS responses are a) lower than expected and b) non-linear should enable the identification of the causes and therefore allow the current ESI-MS method to be developed to ensure linear ESI-MS responses are exhibited by MUPs, for more accurate relative quantification of MUPs in samples of differing complexities, and the possibility of the development of an absolutely quantitative method. It must be remembered, however, that current ESI-MS analysis of intact MUPs is unable to identify some MUP isoforms as the sequence homology of MUPs means the mass spectrometer is unable to resolve MUPs of similar mass (for example, the 18693 and 18694 Da MUPs), suggesting that for absolute quantification of MUPs in urine, peptide-level analysis is required. The use of a multiple reaction monitoring (MRM) method with a triple quadrupole mass spectrometer could improve the quantification of MUPs in urine using QconCAT methodology, providing increased sensitivity and a greater dynamic range for the detection and quantification of less abundant MUP variants in complex sample mixtures. Despite this, the use of the MUP QconCAT faces further limitations arising from the high evolution rate of MUPs. It is possible that new MUP isoforms or allelic variants could contain peptides not represented by the current MUP QconCAT, meaning that further QconCATs would need to be designed.

In Chapter 3, the male specific MUPs have been fully sequenced, and so it would be beneficial to obtain full accurate sequence information for the 18918 Da MUP. The

development of a chromatography method that provides better resolution of MUPs may allow the 18918 Da MUP to be purified from male urine, or employing an alternative mass spectrometry method, such as electron transfer dissociation (ETD), may provide additional sequence coverage. In relation to the behavioural experiments, further experiments need to be undertaken to definitively link male MUP expression and behaviour, as in this thesis, only a relatively small number of animals were sampled. The use of recombinant versions of all four *M. spicilegus* MUPs could then be used in further behavioural experiments, investigating female attraction to the male-specific MUPs, in particular the 18762 Da MUP, and whether any of the MUPs promote aggressive behaviour. This would build significantly on the behavioural experiments conducted in Chapter 3, providing a deeper insight into the unusual social and sexual behaviours this species has been found to exhibit.

The lactation investment from communally nursing female mice was successfully tracked in their pups confirming that no females, related or unrelated, discriminated between their own pups and their female partners'; familiar females cooperatively invest in communal litters to increase reproductive success. In most cases, however, one female invests significantly more in the entire communal litter than the other. For this reason, further communal nursing experiments would be required to discover whether there is a link between investment from a mother and the age of her pups – in more than half of the communal nests, the mother of the older litter invested more, so this is a possible link which is best explored with a greater sample size. To determine whether the differences in investment are observed throughout the communal nursing experiment, or whether the apparent differences in investment are due to mothers taking turns, pup samples would be needed to be removed from the communal nest at intervals during the experiment as well as on the final day.

The continuing work on MUP quantification, focusing on the role of MUPs in chemical communication, is important for pest control strategies in developing countries. Research into animal welfare will also benefit from ongoing MUP quantification studies, as well as the work on the roles of *M. spicilegus* MUPs in chemical communication amongst conspecifics. Animal welfare projects monitor the wellbeing of laboratory rodents, and so MUP quantification will aid the identification of the MUP patterns that trigger behavioural responses such as aggression, whereas the continuing investigation into *M. spicilegus* MUPs will identify whether any MUPs have a role in aggressive behaviour amongst laboratory-kept animals of this species. Continuing the *M. spicilegus* studies should also provide an insight into the of role

scent communication in mating preferences; when these mice first arrived at the Mammalian Behaviour and Evolution group, difficulties arose trying to get them to successfully reproduce. As discussed in Chapter 3, *M. spicilegus* are not cooperative breeders, and so a greater understanding of the biochemical basis of their mating behaviour will improve reproductive success in mice kept in behavioural research laboratories. The study into the communally nursing BALB/c laboratory has also provided an insight into the social and reproductive behaviour, which will further benefit laboratories involved in behavioural research and assist in successful breeding. Many animal welfare projects are based on information generated from experiments in behavioural laboratories, and complementing this knowledge with data from biochemistry experiments will enhance the understanding of complex behaviour in rodents.

Chapter 7: References

Altschul, S. F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool, *J. Mol. Biol*., **215**, p. 403 - 410.

Apanius, V. P., D. Slev, P. Ruff, L. R. Potts, W. K. (1997) The nature of selection on the major histocompatibility complex, *Crit. Rev. Immunol*., **17**, p. 179 - 224.

Armstrong, S. D., Robertson, D. H., Cheetham, S. A., Hurst, J. L., Beynon, R. J. (2005) Structural and functional differences in isoforms of mouse major urinary proteins: a male-specific protein that preferentially binds a male pheromone: *Biochem. J*., **391**, p. 343 - 350.

Bacchini, A., Gaetani, E., Cavaggioni, A. (1992) Pheromone binding proteins of the mouse, *Mus musculus Experientia,* **48**, p. 419 - 421.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review: *Anal. Bioanal. Chem*., **389**, p. 1017-31.

Barr, J. R., Maggio, V.L., Patterson, D.G, Jr., Cooper, G.R., Henderson, L.O., Turner, W.E., Smith, S.J., Hannon, W.H., Needham, L.L., Sampson, E.J. (1996) Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I, *Clin. Chem.,* **42**, p. 1676 - 1682.

Beauchamp, G. K., Yamazaki, K. (2003) Chemical signaling in mice, *Biochem. Soc. Trans*., **31**, p. 147 - 151.

Beynon, R. J., Armstrong, S. D., Claydon, A. J., Davidson, A. J., Eyers, C. E., Langridge, J. I., Gómez-Baena, G., Harman, V. M., Hurst, J. L., Lee, V., McLean, L., Pattison, R., Roberts, S. A., Simpson, D. M., Unsworth, J., Vonderach, M., Williams, J. P., Woolerton, Y. E. (2015) Mass spectrometry for structural analysis and quantification of the Major Urinary Proteins of the

house mouse, *Int. J. Mass Spectrom*., In press.

Beynon, R. J., Armstrong, S. D., Gómez-Baena, G., Lee, V., Simpson, D., Unsworth, J., Hurst, J. L. (2014) The complexity of protein semiochemistry in mammals, *Biochem. Soc. Trans.*, **42**, p. 837 - 845.

Beynon, R. J., Doherty, M.K., Pratt, J.M., Gaskell, S.J. (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides, *Nat. Methods*, **2**, p. 587 - 589.

Beynon, R. J., Hurst, J.L. (2003) Multiple roles of major urinary proteins in the house mouse, *Mus domesticus, Biochem. Soc. Trans*., **31**, p. 142 - 146.

Beynon, R. J., Hurst, J.L (2004) Urinary proteins and the modulation of chemical scents in mice and rats, *Peptides*, **25**, p. 1553-1563.

Beynon, R. J., Hurst, J.L., Turton, M.J., Robertson, D.H.L., Armstrong, S.D., Cheetham, S.A., Simpson, D., MacNicoll, A., Humphries, R.E. (2008) Urinary lipocalins in rodentia: is there a generic model?, *Chem. Signal*, **11**, p. 37-49.

Beynon, R. J., Robertson, D. H. L., Hubbard, S. J., Gaskell, S. J., Hurst, J. L. (1999) The role of protein binding in chemical communication: major urinary proteins in the house mouse, *in* Advances in Chemical Communication in Vertebrates, New York: Plenum. (Ed. by R. E. Johnston), p. 137 - 147.

Beynon, R. J., Veggerby, C., Payne, C. E., Robertson, D. H. L., Gaskell, S. J., Humphries, R. E., Hurst, J. L. (2002) Polymorphism in major urinary proteins: molecular heterogeneity in a wild mouse population, *J. Chem. Ecol*., **28**, p. 1429 - 1446.

Bocskei, Z., Findlay, J.B., North, A.C., Phillips, S.E., Somers, W.S., Wright, C.E., Lionetti, C., Tirindelli, R., Cavaggioni, A. (1991) Crystallization of and preliminary X-ray data for the mouse major urinary protein and rat alpha-2u globulin, *J. Mol. Biol*, **218**, p. 699 - 701.

Bocskei, Z., Groom, C.R., Flower, D.R., Wright, C.E., Phillips, S.E., Cavaggioni, A., Findlay, J.B., North, A.C. (1992) Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography, *Nature*, **360**, p. 186 - 188.

Boehm, A. M., Pütz, S., Altenhöfer, D., Sickmann, A., Falk, M. (2007) Precise protein quantification based on peptide quantification using iTRAQ™, *BMC Bioinformatics*, **8**, p. 214.

Boehm, T., Zufall, F. (2006) MHC peptides and the sensory evaluation of phenotype, *Trends Neurosci*., **29**, p. 100 - 107.

Bonhomme, F. (1992) Genetic diversity and evolution in the genus *Mus, Techniques in the Behavioral and Neural Sciences,* **8**, p. 41 - 56.

Brennan, P. A., Kendrick, K.M. (2006) Mammalian social odours: attraction and individual recognition, *Philos. Trans. R. Soc. Lond. B: Biol. Sci*., **361**, p. 2061 - 2078.

Brennan, P. A., Zufall, F. (2006) Pheromonal communication in vertebrates, *Nature*, **444**, p. 308 - 315.

Brown, J. L., Eklund, A. (1994) A molecular theory of kin recognition: an integrative review, *Am. Nat*., **143**, p. 170 - 196.

Brown, R. E., MacDonald, D.W. (1985) Social Odours in Mammals Vol. 1, New York, *Oxford University Press*.

Bruce, H. M. (1959) An exteroceptive block to pregnancy in the mouse, *Nature*, **184**, p. 105.

Bruce, H. M. (1960) A block to pregnancy in the mouse caused by proximity of strange males, *J. Reprod. Fertil*., **1**, p. 96 - 103.

Bruce, H. M. (1970) Pheromones, *Br. Med. Bull*., **26**, p. 10 - 13.

Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., Garcin, J. (2007) Isotope-labeled protein standards toward absolute quantitative proteomics, *Mol. Cell. Proteomics*, **6**, p. 2139 - 2149.

Brun, V., Masselon, C., Garin, J., Dupuis, A. (2009) Isotope dilution strategies for absolute quantitative proteomics, *J. Proteomics*, **72**, p. 740 - 749.

Buck, L. B. (2000) The molecular architecture of odor and pheromone sensing in mammals, *Cell*, **100**, p. 611 - 618.

Carroll, L. S., Penn, D. J., Potts, W. K. (2002) Discrimination of MHC-derived odors by untrained mice is consistent with divergence in peptide-binding region residues, *Proc. Natl. Acad. Sci. U. S. A*., **19**, p. 2187 - 2192.

Cech, N. B., Enke, C. G. (2001) Practical implications of some recent studies in electrospray ionization fundamentals, *Mass Spectrom. Rev*., **20**, p. 362 – 387.

Cech, N. B., Krone, J.R., Enke, C.G. (2001) Predicting electrospray response from chromatographic retention time, *Anal. Chem*., **73**, p. 208 - 213.

Chamero, P., Marton, T.F., Logan, D.W., Flanagan, K., Cruz, J.R., Saghatelian, A., Cravatt, B.F., Stowers, L. (2007) Identification of protein pheromones that promote aggressive behaviour, *Nature*, **450**, p. 899 - 902.

Cheetham, S. A., Smith, A. L., Armstrong, S. D., Beynon, R. J., Hurst, J. L. (2009) Limited variation in the major urinary proteins of laboratory mice, *Physiol. Behav*., **96**, p. 253-61.

Cheetham, S. A., Thom, M. D., Jury, F., Ollier, W. E. R., Beynon, R. J., Hurst, J. L. (2007) The genetic basis of individual-recognition signals in the mouse, *Curr. Biol*., **17**, p. 1771 – 1777.

Claydon, A. J., Beynon, R. J. (2011) Protein turnover methods in single-celled organisms: dynamic SILAC, *Methods Mol. Biol*., **759**, p. 179 - 195.

Claydon, A. J., Thom, M. D., Hurst, J. L., Beynon, R. J. (2012) Protein turnover: measurement of proteome dynamics by whole animal metabolic labelling with stable isotope labelled amino acids, *Proteomics*, **12**, p. 1194 - 1206.

Colby, D. R., Vandenberg, J.G. (1974) Regulatory effects of urinary pheromones on puberty in the mouse, *Biol. Reprod*., **11**, p. 268 - 279.

Cottrell, J. C., Green, B. N. (1993) Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics, *San Francisco, CA*., p. 146.

Cottrell, J. C., Green, B.N. (1998) MaxEnt: An essential maximum entropy based tool for interpreting multiply-charged electrospray data, *Micromass UK Ltd.,* Application note 212.

Darwish Marie, A., Veggerby, C., Robertson, D.H., Gaskell, S.J., Hubbard, S.J., Martinsen, L., Hurst, J.L., Beynon, R.J. (2001) Effect of polymorphisms on ligand binding by mouse major urinary proteins, *Protein Sci*., **10**, p. 411 - 417.

Del Punta, K., Leinders-Zufall, T., Rodriguez, I., Jukam, D., Wysocki, C.J., Ogawa, S., Zufall, F., Mombaerts, P. (2002) Deficient pheromone responses in mice lacking a cluster of vomeronasal receptor genes, *Nature*, **419**, p. 70 - 74.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist, *Nucleic Acids Res*., **36**, p. 465 - 469.

Doherty, M. K., Whitehead, C., McCormack, H., Gaskell, S. J., Beynon, R. J. (2005) Proteome dynamics in complex organisms: Using stable isotopes to monitor individual protein turnover rates, *Proteomics*, **5**, p. 522 - 533.

Doty, R. L. (1986) Odor-guided behavior in mammals, *Experientia*, **42**, p. 257 - 271.

Doving, K. B., Trotier, D. (1998) Structure and function of the vomeronasal organ, *J.*

Exp. Biol., **201**, p. 2913 - 2925.

Drickamer, L. C., Hoover, J.E. (1979) Effects of urine from pregnant and lactating female house mice on sexual maturation of juvenile females, *Dev. Psychobiol*., **12**, p. 545 - 551.

Dulac, C., Axel, R. (1995) A novel family of genes encoding putative pheromone receptors in mammals, *Cell*, **83**, p. 195 - 206.

Dupuis, A., Hennekinne, J.A., Garin, J., Brun, V. (2008) Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks, *Proteomics*, **8**, p. 4633 - 4636.

Edwards, S. V., Hedrick, P. W. (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection, *Trends Ecol. Evol*., **13**, p. 305 - 311.

Emlen, S. T. (1995) An evolutionary theory of the family, *P. Natl. Acd. Sci. U.S.A*, **92**, p. 8092 - 8099.

Evershed, R. P., Robertson, D. H. L., Beynon, R. J., Green, B. N. (1993) Application of electrospray ionization mass spectrometry with maximum-entropy analysis to allelic 'fingerprinting' of major urinary proteins, *Rapid Commun. Mass Spectrom*., **7**, p. 882 – 886.

Fenn, J. B. (1993) Ion Formation from Charged Droplets—Roles of Geometry, Energy, and Time, *J. Am. Soc. Mass Spectrom.*, **4** p. 524 - 535.

Ferrige, A. G., Seddon, M. J., Green, B. N., Jarvis, S. A., Skilling, J. and Staunton, J. (1992) Disentangling electrospray spectra with maximum entropy, *Rapid Commun. Mass Spectrom*., **6**, p. 707 - 711.

Ferrige, A. G., Seddon, M. J., Jarvis, S., Skilling, J. and Aplin, R. (1991) Maximum entropy deconvolution in electrospray mass spectrometry, *Rapid Commun. Mass Spectrom*., **5**, p. 374 - 377.

Finlayson, J. S., Asofsky, R., Potter, M., Runner, C.C. (1965) Major urinary protein complex of normal mice: origin, *Science*, **149**, p. 981 - 982.

Flower, D. R. (1996) The lipocalin protein family: structure and function, *Biochem. J.*, **318**, p. 1 - 14.

Flower, D. R., North, A.C., Attwood, T.K. (1993) Structure and sequence relationships in the lipocalins and related proteins, *Protein Sci.*, **2**, p. 753 - 761.

Gittleman, J. L. (1985) Functions of communal care in mammals, *Cambridge University Press*, p. 187 - 205.

Gosling, L. M. (1982) A reassessment of the function of scent marking in territories, *Z. Tierpsychol.*, **60**, p. 89 - 118.

Gosling, L. M., Roberts, S.C., Peter, J.B., Slater, J.S.R.C.T.S., Timothy, J.R. (2001) Scent-marking by male mammals: Cheat-proof signals to competitors and mates, *Adv. Stud. Behav.,* Academic Press, p. 169 – 217.

Gosling, L. M., Roberts,S. C., Thornton, E. A., Andrew, M. J. (2000) Life history costs of olfactory status signalling in mice, *Behav. Ecol. Sociobiol.*, **48**, p. 328 - 332.

Gouat, P., Feron, C., Demouron, S. (2003) Seasonal reproduction and delayed sexual maturity in mound-building mice *Mus spicilegus*, *Reprod. Fertil. Dev.*, **15**, p. 187-95.

Grandori, R. (2003) Origin of the conformation dependence of protein charge-state distributions in electrospray ionization mass spectrometry, *J. Mass Spectrom.*, **38** p. 11 - 15.

Green, J. P., Holmes, A. M., Davidson, A. J., Paterson, S., Stockley, P., Beynon, R. J., Hurst, J. L. (2015) The genetic basis of kin recognition in a cooperatively breeding mammal, *Curr. Biol.*, **25**, p. 2631 – 2641.

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nat. Biotechnol*., **17**, p. 994 - 999.

Halpern, M., Martínez-Marcos, A. (2003) Structure and function of the vomeronasal system: an update, *Prog. Neurobiol*., **70**, p. 245 - 318.

Hayes, L. D. (2000) To nest communally or not to nest communally: a review of rodent communal nesting and nursing, *Anim. Behav*., **59**, p. 677-688.

Heck, A. J. a. V. D. H., R.H. (2004) Investigation of intact protein complexes by mass spectrometry *Mass Spectrom. Rev*., **23**, p. 368 - 389.

Herrada, G., Dulac, C. (1997) A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution, *Cell*, **90**, p. 763 - 773.

Ho, C., Lam, C., Chan, M., Cheung, R., Law, L., Lit, L., Ng, K.F., Suen, M.W.M, Tai, H.L. (2003) Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications, *The Clinical Biochemist Reviews,* **24**, p. 3 - 12.

Humphries, R. E., Robertson, D. H. L., Beynon, R. J., Hurst, J.L. (1999) Unravelling the chemical basis of competitive scent marking in house mice, *Anim. Behav*., **58**, p. 1177 - 1190.

Hurst, J. L. (1990 (a)) Urine marking in populations of wild house mice *Mus domesticus* Rutty. I. Communication between males, *Anim. Behav*., **40**, p. 209 - 222.

Hurst, J. L. (1990 (b)) Urine marking in populations of wild house mice Mus domesticus Rutty. III. Communication between the sexes, *Anim. Behav*., **40**, p. 233 - 243.

Hurst, J. L. (1993) The priming effects of urine substrate marks on interaction between

male house mice, *Mus domesticus, Anim. Behav*., **45**, p. 55 - 81.

Hurst, J. L., Beynon, R. J. (2004) Scent wars: the chemobiology of competitive signalling in mice, *Bioessays*, **26**, p. 1288 - 1298.

Hurst, J. L., Payne, C.E., Nevison, C.M., Marie, A.D., Humphries, R.E., Robertson, D.H., Cavaggioni, A., Beynon, R.J. (2001) Individual recognition in mice mediated by major urinary proteins, *Nature*, **414**, p. 631 - 634.

Hurst, J. L., Robertson, D.H.L., Tolladay, U., Beynon, R.J. (1998) Proteins in urine scent marks of male house mice extend the longevity of olfactory signals, *Anim. Behav*., **55**, p. 1289 - 1297.

Hurst, J. L., Thom, M. D., Nevison, C. M., Humphries, R. E., Beynon, R. J. (2005) MHC odours are not required or sufficient for recognition of individual scent owners, *Proc. R. Soc. B*., **272**, p. 715 – 724.

Ilmonen, P., Penn, D.J., Damjanovich, K., Morrison, L., Ghotbi, L., Potts, W.K. (2007) Major Histocompatibility Complex Heterozygosity Reduces Fitness in Experimentally Infected Mice, *Genetics*, **176**, p. 2501 - 2508.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., Mann, M. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein, *Mol. Cell. Proteomics*, **4**, p. 1265 - 1272.

Jemiolo, B., Alberts, J., Sochinski-Wiggins, Harvey, S., Novotny, M. (1985) Behavioural and endocrine responses of female mice to synthetic analogues of volatile compounds in male urine, *Anim. Behav*., **33**, p. 1114 - 1118.

Jemiolo, B., Andreolini, F., Xie, T.M., Wiesler, D., Novotny, M. (1989) Puberty affecting synthetic analogs of urinary chemosignals in the house mouse, *Mus domesticus, Physiol. Behav*., **46**, p. 293 - 298.

Jemiolo, B., Harvey, S., Novotny, M. (1986) Promotion of the Whitten effect in female mice by synthetic analogs of male urinary constituents, *Proc. Natl. Acad. Sci. U.S.A*, **83**, p. 4576 - 4579.

Jemiolo, B., Novotny, M. (1994) Inhibition of sexual maturation in juvenile female and male mice by a chemosignal of female origin, *Physiol. Behav*., **55** p. 519 - 522.

Jemiolo, B., Xie, T. M., Novotny, M. (1992) Urine marking in male mice: responses to natural and synthetic chemosignals, *Physiol. Behav*., **52**, p. 521 - 526.

Jemiolo, B., Xie, T.M., Novotny, M. (1991) Socio-sexual olfactory preference in female mice: attractiveness of synthetic chemosignals, *Physiol. Behav*., **50**, p. 1119 - 1122.

Jennisons, M. D., MacDonald, D. W. (1994) Cooperative breeding in mammals, *Trends Ecol. Evol*., **9**, p. 89 - 93.

Johnson, D., al-Shawi, R., Bishop, J.O. (1995) Sexual dimorphism and growth hormone induction of murine pheromone-binding proteins, *J. Mol. Endocrinol*., **14**, p. 21 - 34.

Johnson, R. P. (1973) Scent marking in mammals, *Anim. Behav*., **21**, p. 521 - 535.

Kaiser, S. E., Riley, B.E., Shaler, T.A., Trevino, R.S., Becker, C.H., Schulman, H., Kopito, R.R. (2011) Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools, *Nat. Methods*, **8**, p. 691 - 696.

Karlson, P., Luscher, M. (1959) 'Pheromones': a new term for a class of biologically active substances, *Nature*, **183**, p. 55 - 56.

Keller, M., Douhard, Q., Baum, M.J., Bakker, J. (2006) Sexual experience does not compensate for the disruptive effects of zinc sulfate—lesioning of the main olfactory epithelium on sexual behavior in male mice, *Chem. Senses*, **31**, p. 753 - 762.

Keller, M., Pierman, S., Douhard, Q., Baum, M.J., Bakker, J. (2006) The vomeronasal organ is required for the expression of lordosis behaviour, but not sex discrimination in female mice, *Eur. J. Neurosci*., **23**, p. 521 - 530.

Kelliher, K. R., Spehr, M., Li, X.H., Zufall, F., Leinders-Zufall, T. (2006) Pheromonal recognition memory induced by TRPC2-independent vomeronasal sensing, *Eur. J. Neurosci.*, **23**, p. 3385 - 3390.

Keverne, E. B. (1983) Pheromonal influences on the endocrine regulation of reproduction, *Trends Neurosci*., **6**, p. 381.

Keverne, E. B. (1999) The vomeronasal organ, *Science*, **286**, p. 716 - 720.

Kimoto, H., Haga, S., Sato, K., Touhara, K. (2005) Sex-specific peptides from exocrine glands stimulate mouse vomeronasal sensory neurons, *Nature*, **437**, p. 898 - 891.

Kimoto, H., Sato, K., Nodari, F., Haga, S., Holy, T.E., Touhara, K. (2007) Sex- and strain-specific expression and vomeronasal activity of mouse ESP family peptides, *Curr. Biol*., **17**, p. 1879 - 1884.

Klein, J. (1986) Natural History of the Histocompatibility Complex, New York, *Wiley*.

Knopf, J. L., Gallagher, J.F., Held, W.A. (1983) Differential, multihormonal regulation of the mouse major urinary protein gene family in the liver, *Mol. Cell. Biol*., **3**, p. 2232 - 2240.

Konig, B. (2006) Non-offspring nursing in mammals: General implications from a case study on house mice, *Cooperation in Primates and Humans. Mechanisms and Evolution*, p. 191-205.

Konig, B., Lindholm, A. K. (2012) The complex social environment of female house mice *(Mus domesticus)*, *Evolution of the House Mouse,* Cambridge University Press, p. 114 – 134.

Konig, B., Riester, J., Markl, H. (1989) Maternal care in house mice (*Mus musculus*): II. The energy cost of lactation as a function of litter size, *J. Zool*., **216**, p. 195 - 210.

Koyama, S. (2004) Primer effects by conspecific odors in house mice: a new perspective in the study of primer effects on reproductive activities, *Horm. Behav*., **46**, p. 303 - 310.

Krauter, K., Leinwand, L., D'Eustachio, P., Ruddle, F., Darnell, J.E. Jr. (1982) Structural genes of the mouse major urinary protein are on chromosome 4, *J. Cell Biol.*, **94**, p. 414 - 416.

Krusemark, C. J., Frey, B.L., Belshaw, P.J., Smith, L.M. (2009) Modifying the charge state distribution of proteins in electrospray ionization mass spectrometry by chemical derivatization, *J Am Soc Mass Spectrom.*, **20**, p. 1617 - 1625.

Kurtz, J., Wegner, K.M., Kalbe, M., Reusch, T.B., Schaschl, H., Hasselquist, D., Milinski, M. (2006) MHC genes and oxidative stress in sticklebacks: an immuno-ecological approach, *Proc. Biol. Sci*., **273**, p. 1407 - 1414.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, **227**, p. 680 - 685.

Lange, V., Picotti, P., Domon, B., Abersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial, *Molecular Systems Biology*, **4**, p. 222.

Leinders-Zufall, T., Brennan, P., Widmayer, P., S, P.C., Maul-Pavicic, A., Jager, M., Li, X.H., Breer, H., Zufall, F., Boehm, T. (2004) MHC class I peptides as chemosensory signals in the vomeronasal organ, *Science*, **306**, p. 1033 - 1037.

Leinders-Zufall, T., Lane, A.P., Puche, A.C., Ma, W., Novotny, M.V., Shipley, M.T., Zufall, F. (2000) Ultrasensitive pheromone detection by mammalian vomeronasal neurons, *Nature*, **405**, p. 792 - 796.

Liberles, S. D., Buck, L.B. (2006) A second class of chemosensory receptors in the olfactory epithelium, *Nature*, **442**, p. 645 - 650.

Lin, D. Y., Zhang, S-Z., Block, E., Katz, L.C. (2005) Encoding social signals in the mouse main olfactory bulb, *Nature*, **434**, p. 470 - 477.

Lin, W., Margolskee, R., Donnert, G., Hell, S.W., Restrepo, D. (2007) Olfactory neurons expressing transient receptor potential channel M5 (TRPM5) are involved in sensing semiochemicals, *Proc. Natl. Acad. Sci. U.S.A*, **104**, p. 2471 - 2476.

Liu, H., Sadygov, R.G., Yates, J.R. (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics, *Anal. Chem*., **76**, p. 4193 - 4201.

Logan, D. W., Marton, T.F., Stowers, L. (2008) Species specificity in major urinary proteins by parallel evolution, *PLoS ONE*, **3**, p. e3280.

Lucke, C., Franzoni, L., Abbate, F., Lohr, F., Ferrari, E., Sorbi, R.T., Ruterjans, H., Spisni, A. (1999) Solution structure of a recombinant mouse major urinary protein, *Eur. J. Biochem*, **266**, p. 1210 - 1218.

Ma, M. (2010) Multiple Olfactory Subsystems Convey Various Sensory Signals, *In: The Neurobiology of Olfaction*, CRC Press, Chaper 9.

Ma, W., Miao, Z., Novotny, M.V. (1998) Role of the adrenal gland and adrenal mediated chemosignals in suppression of estrus in the house mouse: the leeboot effect revisited, *Biol. Reprod*., **59**, p. 1317 - 1320.

Ma, W., Miao, Z., Novotny, M.V. (1999) Induction of estrus in grouped female mice (*Mus domesticus*) by synthetic analogues of preputial gland constituents, *Chem. Senses*, **24**, p. 289 - 293.

Macrides, F., Bartke, A., Dalterio, S. (1975) Strange females increase plasma

testosterone levels in male mice, *Science*, **189**, p. 1104 - 1106.

Manning, C. J., Dewsbury, D. A., Wakeland, E. K., Potts, W. K. (1995) Communal nesting and communal nursing in house mice,*Mus musculus domesticus, Anim. Behav.*, **50**, p. 741 - 751.

Matsunami, H., Buck, L.B. (1997) A multigene family encoding a diverse array of putative pheromone receptors in mammals, *Cell*, **90**, p. 775 - 784.

Meredith, M. (1994) Chronic recording of vomeronasal pump activation in awake behaving hamsters, *Physiol. Behav*., **56**, p. 345 - 354.

Mudge, J. M., Armstrong, S.D., McLaren, K., Beynon, R.J., Hurst, J.L., Nicholson, C., Robertson, D.H., Wilming, L.G., Harrow, J.L. (2008) Dynamic instability of the major urinary protein gene family revealed by genomic and phenotypic comparisons between C57 and 129 strain mice, *Genome Biol*., **9**, p. R91.

Mugford, R. A., Nowell, N.W. (1971) The relationship between endocrine status of female opponents and aggressive behaviour of male mice, *Anim. Behav*., **19**, p. 153 - 155.

Mugicnat-Caretta, C., Cavaggioni, A., Caretta, A. (2004) Male urinary chemosignals differentially affect aggressive behaviour in male mice, *J. Chem. Ecol*., **30**, p. 777 - 791.

Nevison, C. M., Armstrong, S., Beynon, R.J., Humphries, R.E., Hurst, J.L. (2003) The ownership signature in mouse scent marks is involatile, *Proc. Biol. Sci*., **270**, p. 1957 - 1963.

Novotny, M., Harvey, S., Jemiolo, B. (1990) Chemistry of male dominance in the house mouse, *Mus domesticus*, *Experientia*, **46**, p. 109 - 113.

Novotny, M., Harvey, S., Jemiolo, B., Alberts, J. (1985) Synthetic pheromones that promote inter-male aggression in mice, *Proc. Natl. Acad. Sci. U.S.A*., **82**, p.

2059 - 2061.

Novotny, M. V. (2003) Pheromones, binding proteins and receptor responses in rodents, *Biochem. Soc. Trans*., **31**, p. 117 - 122.

Novotny, M. V., Jemiolo, B., Harvey, S., Wiesler, D., Marchlewska-Koj, A. (1986) Adrenal-mediated endogenous metabolites inhibit puberty in female mice, *Science*, **231**, p. 722 - 725.

Novotny, M. V., Jemiolo, B., Wiesler, D., Ma, W., Harvey, S., Xu, F., Xie, T.M., Carmack, M. (1999) A unique urinary constituent, 6-hydroxy-6-methyl-3-heptanone, is a pheromone that accelerates puberty in female mice, *Chem. Biol*., **6**, p. 377 - 383.

Novotny, M. V., Soini, H.A., Koyama, S., Wiesler, D., Bruce, K.E., Penn, D.J. (2007) Chemical Identification of MHC-influenced Volatile Compounds in Mouse Urine. I: Quantitative Proportions of Major Chemosignals, *J. Chem. Ecol*., **33**, p. 417 - 434.

Null, A. P., Nepomuceno, A.I., Muddiman, D.C. (2003) Implications of Hydrophobicity and Free Energy of Solvation for Characterization of Nucleic Acids by Electrospray Ionization Mass Spectrometry, *Anal. Chem*., **75** p. 1331 - 1339.

Nyby, J., Dizinno, G., Whitney, G. (1977) Sexual dimorphism in ultrasonic vocalizations of mice (*Mus musculus*): gonadal hormone regulation, *J. Comp. Physiol. Psychol*., **91**, p. 1424 - 1431.

Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J.R., Resing, K.A., Ahn, N.G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics, *Mol. Cell. Proteomics*, **4**, p. 1487 - 1502.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H, Pandey, A., Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC,

as a simple and accurate approach to expression proteomics, *Mol. Cell. Proteomics*, **1**, p. 376-86.

Osaka, I., Takayama, M. (2014) Influence of hydrophobicity on positive- and negative-ion yields of peptides in electrospray ionization mass spectrometry, *Rapid Commun Mass Spectrom*., **28**, p. 2222 - 2226.

Packer, C., Lewis, S., Pusey, A. (1992) A comparative analysis of non-offspring nursing, *Anim. Behav.*, **43**, p. 265 - 281.

Patris, B., Baudoin, C. (1998) Female sexual preferences differ in *Mus spicilegus* and *Mus musculus domesticus*: the role of familiarization and sexual experience, *Anim. Behav*., **56**, p. 1465 - 1470.

Patris, B., Baudoin, C. (2000) A comparative study of parental care between two rodent species: implications for the mating system of the mound-building mouse *Mus spicilegus*, *Behav. Processes,* **51**, p. 35-43.

Patris, B., Gouat, P., Jacquot, C., Christophe, N., Baudoin, C. (2002) Agonistic and sociable behaviors in the mound-building mice, *Mus spicilegus*: a comparative study with *Mus musculus domesticus*, *Aggressive behav*., **28**, p. 75-84.

Penn, D. J. (2002) Major Histocompatability Complex (MHC), *Encyclopedia of Life Sciences*, Macmillan Publishers Ltd, Nature Publishing Group.

Penn, D. J., Potts, W. K. (1998 (a)) How do major histocompatibility complex genes influence odor and mating preferences?, *Adv. Immunol*., **69**, p. 411 - 436.

Penn, D. J., Potts, W. K. (1998 (b)) Untrained mice distinguish MHC-determined odors, *Physiol. Behav.*, **64**, p. 235 - 243.

Penn, D. J., Potts, W. K. (1999) The evolution of mating preferences and major histocompatibility complex genes, *Am. Nat*., **153**, p. 145 - 164.

Perkins, D. N., Pappin, D.J., Creasy, D.M., Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data, *Electrophoresis*, **20**, p. 3551 - 3567.

Petrulis, A. (2013) Chemosignals, Hormones and Mammalian Reproduction, *Horm. Behav*., **63**, p. 723 - 741.

Phelan, M. M., McLean, L., Armstrong, S. D., Hurst, J. L., Beynon, R. J., Lian, L. (2014) The structure, stability and pheromone binding of the male mouse sex pheromone darcin, *PLoS ONE*, **9**, p. 1 – 16.

Pitt, J. J. (2009) Principles and Applications of Liquid Chromatography-Mass Spectrometry in Clinical Biochemistry, *Clin. Biochem. Rev.,* **30**, p. 19 - 34.

Potts, W. K., Wakeland, E.K. (1993) Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference, *Trends Genet*., **9**, p. 408 - 412.

Pratt, J. M., Simpson, D.M., Doherty, M.K., Rivers, J., Gaskell, S.J., Beynon, R.J. (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes, *Nat. Protoc*., **1**, p. 1029 - 1043.

Ramm, S. A., Cheetham, S.A., Hurst, J.L. (2008) Encoding choosiness: female attraction requires prior physical contact with individual male scents in mice, *Proc. Biol. Sci.*, **275**, p. 1727 - 1735.

Rappsilber, J., Ryder, U., Lamond, A.I., Mann, M. (2002) Large-scale proteomic analysis of the human spliceosome, *Genome Res.*, **12**, p. 1231 - 1245.

Restrepo, D., Arellano, J., Oliva, A.M., Schaefer, M.L., Lin, W.H. (2004) Emerging views on the distinct but related roles of the main and accessory olfactory systems in responsiveness to chemosensory signals in mice, *Horm. Behav*., **46**, p. 247 - 256.

Rich, T. J., Hurst, J. L. (1998) Scent marks as reliable signals of the competitive ability of mates, *Anim. Behav*., **56**, p. 727 - 735

Rich, T. J., Hurst, J.L. (1999) The competing countermarks hypothesis: reliable assessment of competitive ability by potential mates, *Anim. Behav*., **58**, p. 1027 - 1037.

Rivers, J., McDonald, L., Edwards, I. J., Beynon, R. J. (2008) Asparagine deamidation and the role of higher order protein structure, *J. Proteome. Res*., **7**, p. 921-7.

Roberts, S. A., Davidson, A. J., McLean, L., Beynon, R. J., Hurst, J. L. (2012) Pheromonal induction of spatial learning in mice, *Science*, **338**, p. 1462-5.

Roberts, S. A., Simpson, D. M., Armstrong, S. D., Davidson, A. J., Robertson, D. H., McLean, L., Beynon, R. J., Hurst, J. L. (2010) Darcin: a male pheromone that stimulates female memory and sexual attraction to an individual male's odour, *BMC Biol.*, **8**, p. 75.

Robertson, D. H., Cox, K.A., Gaskell, S.J., Evershed, R.P., Beynon, R.J. (1996) Molecular heterogeneity in the Major Urinary Proteins of the house mouse *Mus musculus*, *Biochem. J.*, **316**, p. 265 - 272.

Robertson, D. H., Hurst, J. L., Searle, J. B., Gunduz, I., Beynon, R. J. (2007) Characterization and comparison of major urinary proteins from the house mouse, *Mus musculus domesticus*, and the aboriginal mouse, *Mus macedonicus*, *J. Chem. Ecol.*, **33**, p. 613 - 630.

Robertson, D. H., Hurst, J.L., Bolgar, M.S., Gaskell, S.J., Beynon, R.J. (1997) Molecular heterogeneity of urinary proteins in wild house mouse populations, *Rapid Commun. Mass. Spectrom*. **11**, p. 786 - 790.

Robertson, D. H. L., Beynon, R. J., Evershed, R. P. (1993) Extraction, characterization, and binding analysis of two pheromonally active ligands associated with major urinary protein of house mouse (*Mus*

musculus), *J. Chem. Ecol*., **19**, p. 1405 - 1416.

Robertson, D. H. L., Hurst, J.L., Hubbard, S., Gaskell, S.J., Beynon, R.J. (1998) Ligands of urinary lipocalins from the mouse: Uptake of environmentally derived chemicals, *J. Chem. Ecol*., **24**, p. 1127-1140.

Robertson, D. L., Marie, A., Veggerby, C., Hurst, J., Beynon, R. (2001) Characteristics of Ligand Binding and Release by Major Urinary Proteins, *In: Chemical Signals in Vertebrates 9*, Springer US, p. 169 - 176.

Ruemke, P., Thung, P.J. (1964) Immunological Studies on the Sex-Dependent Prealbumin in Mouse Urine and on Its Occurrence in the Serum, *Acta. Endocrinol. (Copenhagen)*, **47**, p. 156 - 164.

Ryba, N. J., Tirindelli, R. (1997) A new multigene family of putative pheromone receptors, *Neuron,* **19**, p. 371 - 379.

Schoenheimer, R., Rittenberg, D., Foster, G. L., Keston, A. S., Ratner, S. (1938) The application of the nitrogen isotope N15 for the study of protein metabolism, *Science,* **88**, p. 599 - 600.

Schulz, K. M., Molenda-Figueira, H.A., Sisk, C.L. (2009) Back to the future: The organizational-activational hypothesis adapted to puberty and adolescence, *Horm. Behav*., **55**, p. 597 - 604.

Schwende, F. J., Wiesler, D., Novotny, M. (1984) Volatile compounds associated with estrus in mouse urine:potential pheromones, *Experientia*, **40**, p. 213 - 215.

Shahan, K., Gilmartin, M., Derman, E. (1987) Nucleotide sequences of liver, lachrymal, and submaxillary gland mouse major urinary protein mRNAs: mosaic structure and construction of panels of gene-specific synthetic oligonucleotide probes, *Mol. Cell. Biol*., **7**, p. 1938 - 1946.

Sharrow, S. D., Novotny, M.V., Stone, M.J. (2003) Thermodynamic analysis

of binding between mouse major urinary protein-I and the pheromone 2-secbutyl-4,5-dihydrothiazole, *Biochemistry*, **42**, p. 6302 - 6309.

Shaw, P. H., Held, W.A., Hastie, N.D. (1983) The gene family for major urinary proteins: expression in several secretory tissues of the mouse, *Cell*, **32**, p. 755 - 761.

Sheehan, M. J., Lee, V., Corbett-Detig, R., Bi, K., Beynon, R. J., Hurst, J. L., Nachman, M. W. (2015) Selection on coding and regulatory variation among paralogs of the Mup gene family maintains individuality in urinary scent markings of a wild house mouse population, *PLOS Biology*, Under review.

Sherborne, A. L., Thom, M.D., Paterson, S., Jury, F., Ollier, W.E., Stockley, P., Beynon, R.J., Hurst, J.L. (2007) The genetic basis of inbreeding avoidance in house mice, *Curr. Biol*., **17**, p. 2061 - 2066.

Shi, Y., Rodriguez, M., Shahan, K., Derman, E. (1989) Subfamily of submaxillary gland-specific Mup genes: chromosomal linkage and sequence comparison with liver-specific Mup genes, *Nucleic Acids Res*., **17**, p. 6191 - 6203.

Silva, J. C., Gorenstein, M.V., Li, G.Z., Vissers, J.P., Geromanos, S.J. (2006) Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition, *Mol. Cell. Proteomics*, **5**, p. 144 - 156.

Silver, L. M. (1995) Mouse Genetics - Concepts and Applications, *Oxford University Press*.

Singh, P. B., Brown, R.E., Roser, B. (1987) MHC antigens in urine as olfactory recognition cues, *Nature*, **327**, p. 161 - 164.

Soini, H. A., Bruce, K.E., Wiesler, D., David, F., Sandra, P., Novotny, M.V. (2005) Stir bar sorptive extraction: A new quantitative and comprehensive sampling technique for determination of chemical signal profiles from biological media, *J. Chem. Ecol*., **31**, p. 377 - 392.

Sokolov, V. E., Kotenkova, E.V., Michailenko, A.G. (1998) *Mus spicilegus, Mammal spec*., **592**, p. 1 - 6.

Spehr, M., Kelliher, K.R., Li, X.H., Boehm, T., Leinders-Zufall, T., Zufall, F. (2006) Essential role of the main olfactory system in social recognition of major histocompatibility complex peptide ligands, *J. Neurosci*., **26**, p. 1961 - 1970.

Spiegelberg, T., Bishop, J.O. (1988) Tissue-specific gene expression in mouse hepatocytes cultured in growth-restricting medium, *Mol. Cell. Biol*., **8**, p. 3338 - 3344.

Stopka, P., Janotova, K., Heyrovsky, D. (2007) The advertisement role of major urinary proteins in mice, *Physiol. Behav*., **5**, p. 667 – 670.

Stowers, L., Holy, T.E., Meister, M., Dulac, C., Koentges, G. (2002) Loss of sex discrimination and male–male aggression in mice deficient for TRP2, *Science,* **295**, p. 1493 - 1500.

Stöcklin, R., Vu, L., Vadas, L., Cerini, F., Kippen, A.D., Offord, R.E., Rose, K. (1997) A Stable Isotope Dilution Assay for the In Vivo Determination of Insulin Levels in Humans by Mass Spectrometry, *Diabetes*, **46**, p. 45 - 50.

Swaney, W. T., Keverne, E.B. (2009) The evolution of pheromonal communication, *Behav. Brain. Res*., **200**, p. 239 - 247.

Swanson, L. W., Petrovich, G.D. (1998) What is the amygdala?, *Trends Neurosci*., **21**, p. 323 - 331.

Szenczi, P., Bánszegi, O., Groó, Z., Altbäcker, V. (2012) Development of the social behaviour of two mice species with contrasting social systems, *Aggressive Behav*., **38**, p. 288-297.

Thom, M. D., Stockley, P., Jury, F., Ollier, W. E. R., Beynon, R. J., Hurst, J. L. (2008) The direct assessment of genetic heterozygosity through scent in the mouse,

Curr. Biol., **18**, p. 619 – 623.

Thompson, R. N., McMillon, R., Napier, A., Wekesa, K.S. (2007) Pregnancy block by MHC class I peptides is mediated via the production of inositol 1,4,5-trisphosphate in the mouse vomeronasal organ, *J. Exp. Biol*., **210**, p. 1406 - 1412.

Timm, D. E., Baker, L.J., Mueller, H., Zidek, L., Novotny, M.V. (2001) Structural basis of pheromone binding to mouse major urinary protein (MUP-I), *Protein Sci*., **10**, p. 997 - 1004.

Tirindelli, R., Dibattista, M., Pifferi, S., Menini, A. (2009) From Pheromones to Behavior, *Physiol. Rev*., **89**, p. 921 - 956.

Tong, W., Hoekstra, H. (2012) *Mus spicilegus*, *Curr. Biol*., **22**, p. 858 - 859.

Unsworth, J. (2014) Identification, Characterisation and Quantification of Proteins used in Chemical Communication, *Thesis*, University of Liverpool.

Van Der Lee, S., Boot, L.M. (1956) Spontaneous pseudopregnancy in mice,, *Acta. Physiol. Pharmacol. Neer*., **5**, p. 213 - 215.

Vandenbergh, J. G. (1969) Male odor accelerates female sexual maturation in mice, *Endocrinology*, **84**, p. 658 - 660.

Waterlow, J. C., Garlick, P. J., Millward, D. J. (1978) Protein turnover in mammalian tissues and in the whole body, *Elsevier/North Holland Biomedical Press*, Amsterdam, The Netherlands.

Weber, E. M., Olsson, I. A. S. (2008) Maternal behaviour in *Mus musculus* sp.: An ethological review, *Appl. Anim. Behav. Sci*., **114**, p. 1 - 22.

Whitten, W. K., Bronson, F.H., Greenstein, J.A. (1968) Estrus-inducing pheromone of male mice: transport by movement of air, *Science*, **161**, p. 584 - 585.

Wiley, R. H. (1983) The evolution of communication: information and manipulation, *Anim. Behav*., **2**, Blackwell Scientific Publications, p. 157 - 189.

Wilkins, M. R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., Williams, K.L., Hochstrasser, D.F. (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis, *Biotechnology*, **14**, p. 61 - 65.

Wilkinson, G. S., Baker, A.E.M. (1988) Communal nesting among genetically similar house mice, *Ethology*., **77**, p. 103 - 114.

Wyatt, T. D. (2009) Pheromones and Animal Behaviour: Communication by Smell and Taste, *Pragmatics and Cognition*, p. 482 - 490.

Xu, F., Schaefer, M., Kida, I., Schafer, J., Liu, N., Rothman, DL., Hyder, F., Restrepo, D., Shepherd, D.M. (2005) Simultaneous activation of mouse main and accessory olfactory bulbs by odors or pheromones, *J. Comp. Neurol*., **489**, p. 491 - 500.

Yamaguchi, M., Yamazaki, K., Beauchamp, G. K., Bard, J., Thomas, L., Boyse, E. A. (1981) Distinctive urinary odors governed by the major histocompatibility locus of the mouse, *Proc. Natl. Acad. Sci. U.S.A*, **78**, p. 5817 - 5820.

Yamazaki, K., Beauchamp, G. K., Curran, M., Bard, J., Boyse, E. A. (2000) Parent-progeny recognition as a function of MHC odortype identity, *Proc. Natl. Acad. Sci. U. S. A*, **97**, p. 10500 - 105002.

Yamazaki, K., Beauchamp, G. K., Egorov, I. K., Bard, J., Thomas, L., Boyse, E. A. (1983) Sensory distinction between H-2b and H-2bm1 mutant mice, *Proc. Natl. Acad. Sci. U. S. A*, **80**, p. 5685 - 5688.

Yamazaki, K., Boyse, E. A., Mike, V., Thaler, H. T., Mathieson, B. J., Abbott, J., Boyse, J., Zayas, Z. A., Thomas, L. (1976) Control of mating preferences in

mice by genes in the major histocompatibility complex, *J. Exp. Med.*, **144**, p. 1324 - 1335.

Yang, M., Crawley, J.N. (2009) Simple Behavioral Assessment of Mouse Olfaction, *Current Protocols in Neuroscience / Editorial Board, Jacqueline N Crawley*, Unit-8.24.

Zhang, X., Firestein, S. (2002) The olfactory receptor gene superfamily of the mouse, *Nat. Neurosci.*, **5**, p. 124 - 133.

Zhu, W., Smith, J. W., Huang, C. M. (2010) Mass spectrometry-based label-free quantitative proteomics, *J. Biomed. Biotechnol.*, p. 840518.

Zidek, L., Stone, M.J., Lato, S.M., Pagel, M.D., Miao, Z., Ellington, A.D., Novotny, M.V. (1999) NMR mapping of the recombinant mouse major urinary protein I binding site occupied by the pheromone 2-sec-butyl-4,5-dihydrothiazole, *Biochemistry*, **38**, p. 9850 - 9861.

Contributions to publications

- Beynon, R. J., Armstrong, S. D., Claydon, A. J., Davidson, A. J., Eyers, C. E., Langridge, J. I., Gómez-Baena, G., Harman, V. M., Hurst, J. L., **Lee, V**., McLean, L., Pattison, R., Roberts, S. A., Simpson, D. M., Unsworth, J., Vonderach, M., Williams, J. P., Woolerton, Y. E. (2015) Mass spectrometry for structural analysis and quantification of the Major Urinary Proteins of the house mouse, *Int. J. Mass Spectrom*., **391**, p. 146 – 156.

I was responsible for the ESI-MS intact mass analysis of C57BL/6 urine and recombinant MUP mixtures, and provided the data for Figure 6.

- Sheehan, M. J., **Lee, V**., Corbett-Detig, R., Bi, K., Beynon, R. J., Hurst, J. L., Nachman, M. W. (2016) Selection on coding and regulatory variation among paralogs of the Mup gene family maintains individuality in urinary scent markings of a wild house mouse population, *PLOS Genetics*, **12**, e1005891.

I was responsible for the ESI-MS intact mass analysis of mouse urine, alignment of spectra using SpecAlign, and provided input for the writing of the methods section.

- Beynon, R. J., Armstrong, S. D., Gómez-Baena, G., **Lee, V**., Simpson, D., Unsworth, J., Hurst, J. L. (2014) The complexity of protein semiochemistry in mammals, *Biochem. Soc. Trans*., **42**, p. 837 - 845.

I was responsible for the ESI-MS intact mass analysis of C57BL/6 urine and provided the data for Figure 2.

- Application note: Kramer, G., Dekker, N., Gethings, L. A., Shockcor, J. P., **Lee, V**., Beynon, R. J., Langride, J. I., Vissers, J. P. C., Aerts, J. M. F. G. (2014) A Label-Free, Multi-Omic Study to Qualitatively and Quantitatively Characterize the Effects of a Glucosylceramide Inhibitor on Obesity, *Waters Corporation*, Library no. APNT134808146.

I was responsible for the acquisition of protein and lipid LC-MS data, along with some data analysis using Progenesis QI.

Peptide mass (Da)	Peptide sequence	Protease	PEAKS™ confidence score (%)
1566.7	EEASSTGRNFNVEK	Lys-C	88
1595.8	INGEWHTIILASDK	Lys-C	94
1013.6	VLENSLVLK	Lys-C	88
2103.9	VHTVRDEEC*SELSMVADK	Lys-C	75
1138.5	TERAGEYSVK	Lys-C	89
1301.6	YDGFNTFTIPK	Lys-C	82
1856.8	TDYDNFMMSHLINEK	Lys-C	90
2469.2	DGQTFQLMGL_____LSSDIK	Lys-C	75
1179.5	ERFAQLC*EK	Lys-C	96
2591.3	HGILRENIIDLSNANRC*LQARE	Lys-C	69
1438.6	EASSTGRNFNVE	Glu-C	84
1467.8	WHTIILASDKRE	Glu-C	90
1593.8	KIEDNGNFRLFLE	Glu-C	95
746.6	HIHVLE	Glu-C	93
1637.9	NSLVLKVHTVRDEE	Glu-C	94
992.5	LSMVADKTE	Glu-C	95
2273.1	YSVKYDGFNTFTIPKTDYD	Glu-C	82
1234.5	NFMMSHLINE	Glu-C	82
1677.8	GLYGREPDLSSDIKE	Glu-C	71
979.4	RFAQLC*E	Glu-C	92
851.5	KHGILRE	Glu-C	94
1943.9	NIIDLSNANRC*LQARE	Glu-C	74

Supplementary material A: 18742 Da *M. spicilegus* MUP peptides identified by PEAKS *de novo* sequence analysis.
All sequences were assigned a high confidence score by the software. In the sequences, * signifies the post-translational modification carbamidomethylation.

Peptide mass (Da)	Peptide sequence	Protease	PEAKS™ confidence score (%)
1566.7	EEASSTGRNFNVEK	Lys-C	94
1595.8	INGEWHTIILASDK	Lys-C	94
1263.7	IHVLENSLVLK	Lys-C	89
2121.9	FHLGRDEEC*SELSMVADK	Lys-C	76
2008.9	AGEYSVTYDGFNTFTIPK	Lys-C	56
1870.8	TDYDNFMMTHLINEK	Lys-C	96
2846.1	DGETFQLMGLYGREPDLSSDIK	Lys-C	89
1179.5	ERFAQLC*EK	Lys-C	96
1506.9	HGILRENIIDLSK	Lys-C	97
1116.5	ANRC*LQARE	Lys-C	86
1438.6	EASSTGRNFNVE	Glu-C	95
1979.9	EASSTGRNFNVEKINGE	Glu-C	88
1967.0	WHTIILASDKREKIEEH	Glu-C	89
1131.6	HGNFRLFLE	Glu-C	98
737.4	QIHVLE	Glu-C	85
1655.8	NSLVLKFHLGRDEE	Glu-C	92
2032.0	NSLVLKFHLGRDEEC*SE	Glu-C	89
1377.7	LSMVADKTEKAGE	Glu-C	100
1131.5	KAGEYSVTYD	Glu-C	98
1517.7	GFNTFTIPKTDYD	Glu-C	84
1248.6	NFMMTHLINE	Glu-C	96
1677.7	NFMMTHLINEKDGE	Glu-C	95
1525.7	TFQMGLYGREPD	Glu-C	91
922.4	RFAQLC*E	Glu-C	90
851.5	KHGILRE	Glu-C	85
1444.6	LSKANRC*LQARE	Glu-C	75

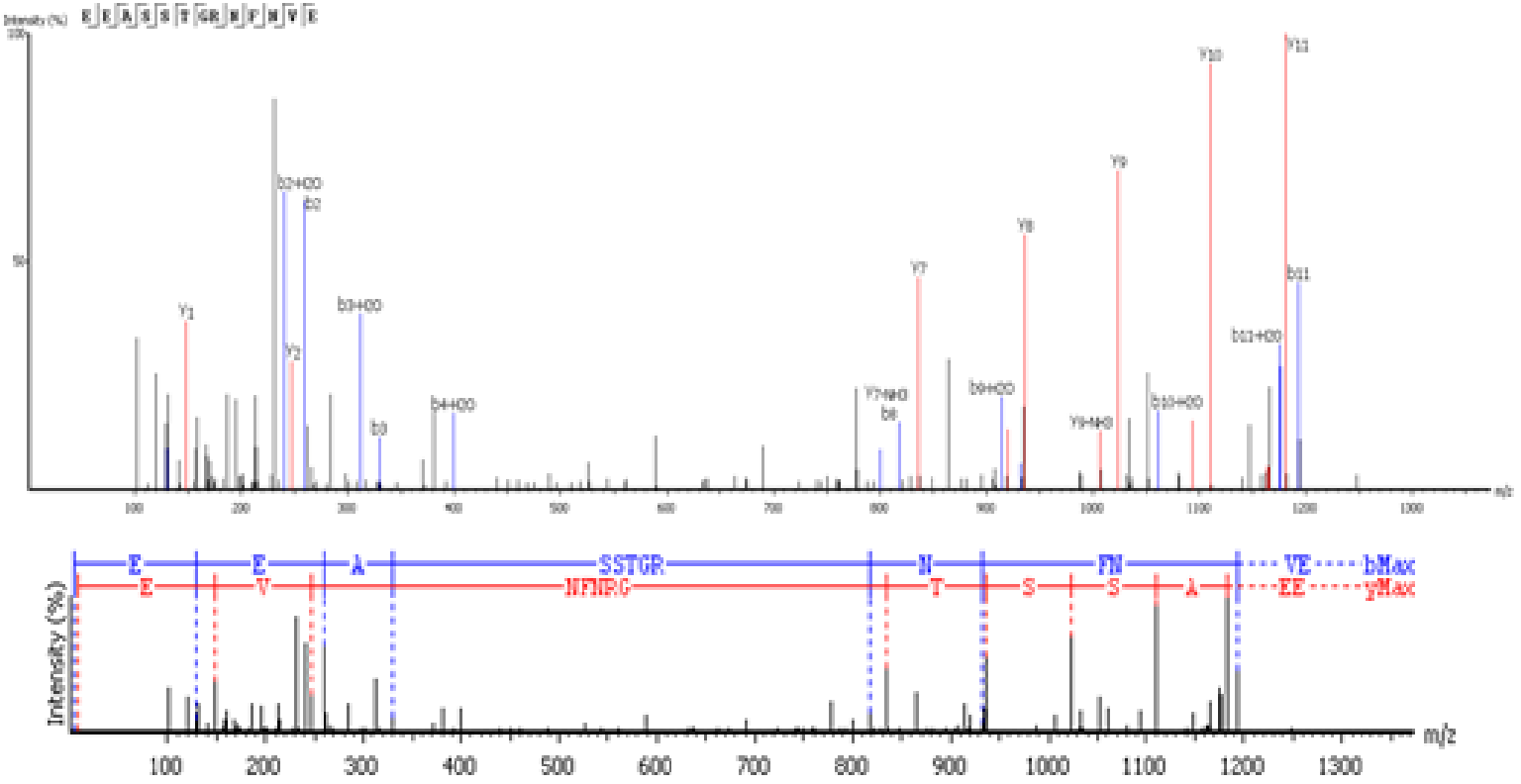
Supplementary material A: 18762 Da *M. spicilegus* MUP peptides identified by PEAKS *de novo* sequence analysis.
All sequences were assigned a high confidence score by the software. In the sequences, * signifies the post-translational modification carbamidomethylation.

Peptide mass (Da)	Peptide sequence	Protease	PEAKS™ confidence score (%)
1566.7	EEASSTGRNFNVEK	Lys-C	88
1595.8	INGEWHTIILASDK	Lys-C	94
801.5	ENSLVLK	Lys-C	83
1947.9	VHTVDEEC*SELSMVADK	Lys-C	96
1138.5	TERAGEYSVK	Lys-C	91
1301.6	YDGFNTFTIPK	Lys-C	91
1856.8	TDYDNFMMSHLINEK	Lys-C	93
2469.2	DGQTFQLMGL_____LSSDIK	Lys-C	75
1179.5	ERFAQLC*EK	Lys-C	96
2591.3	ANRC*LQARE	Lys-C	82
1438.6	EASSTGRNFNVE	Glu-C	84
1979.9	EASSTGRNFNVEKINGE	Glu-C	79
1467.8	WHTIILASDKRE	Glu-C	90
1593.8	KIEDNGNFRLFLE	Glu-C	95
746.6	HIHVLE	Glu-C	93
1481.8	NSLVLKVHTVDEE	Glu-C	98
1857.9	NSLVLKVHTVDEEC*SE	Glu-C	96
992.5	LSMVADKTE	Glu-C	89
1405.7	LSMVADKTERAGE	Glu-C	87
2273.1	YSVKYDGFNTFTIPKTDYD	Glu-C	82
2298.1	TFQLMGLYGREPDLSSDIKE	Glu-C	64
922.4	RFAQLC*E	Glu-C	92
851.5	KHGILRE	Glu-C	94
1886.9	NIIDLSNANRC*LQARE	Glu-C	74

Supplementary material A: 18585 Da *M. spicilegus* MUP peptides identified by PEAKS *de novo* sequence analysis.
All sequences were assigned a high confidence score by the software. In the sequences, * signifies the post-translational modification carbamidomethylation.

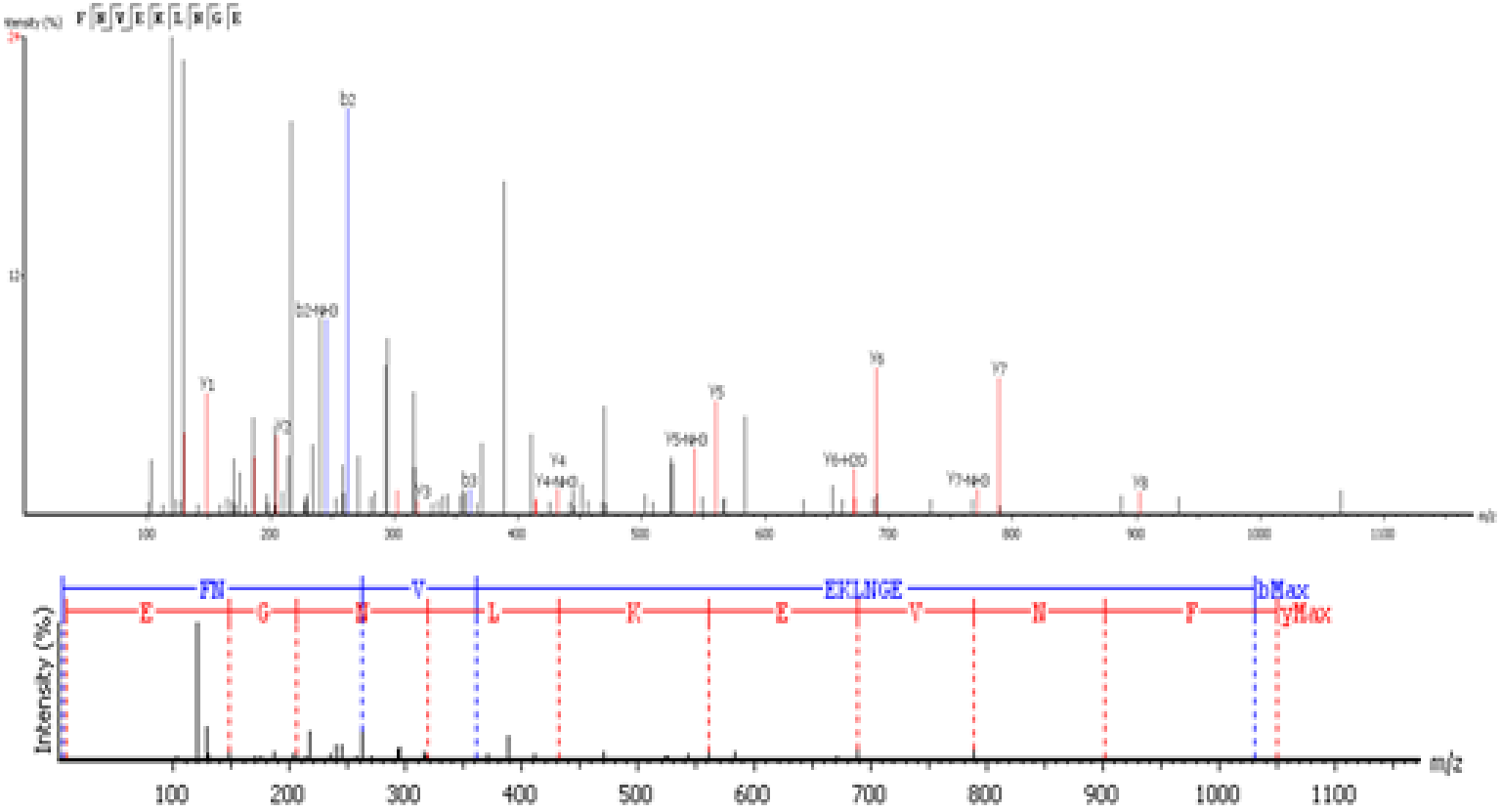
Peptide mass (Da)	Peptide sequence	Protease	PEAKS™ confidence score (%)
1597.7	EEASSMGRNFNVEK	Lys-C	88
1622.8	INGEWYTIILASDK	Lys-C	94
1181.3	C*SEIFLVADK	Lys-C	88
1450.5	AGEYSVTYDGFNK	Lys-C	75
606.7	FTVLK	Lys-C	89
1881.1	TDYDNYIMIHLINKK	Lys-C	82
1606.8	SLYGREPDLNSDLK	Lys-C	76
1663.7	YGREPDLNSDLKEK	Lys-C	50
2039.3	LC*EEHGILRENIIDVTK	Lys-C	75
1179.5	TNRC*LQARE	Lys-C	96
1469.6	EEASSMGRNFNVE	Glu-C	69
1994.3	WYTIILASDKREKIEE	Glu-C	84
1075.3	HGSMRLFVE	Glu-C	95
746.6	HIHVLE	Glu-C	95
1976.1	NSLGFKFKWTDEKC*SE	Glu-C	93
1035.2	IFLVADKTE	Glu-C	94
992.5	IFLVADKTEKAGE	Glu-C	95
817.4	LNSDIKE	Glu-C	79
2273.1	KFVKLC*EE	Glu-C	82
1234.5	HGILRE	Glu-C	82
1677.8	NIIDVTKTNRC*LQARE	Glu-C	71

Supplementary material A: 18918 Da *M. spicilegus* MUP peptides identified by PEAKS *de novo* sequence analysis.
All sequences were assigned a high confidence score by the software. In the sequences, * signifies the post-translational modification carbamidomethylation.



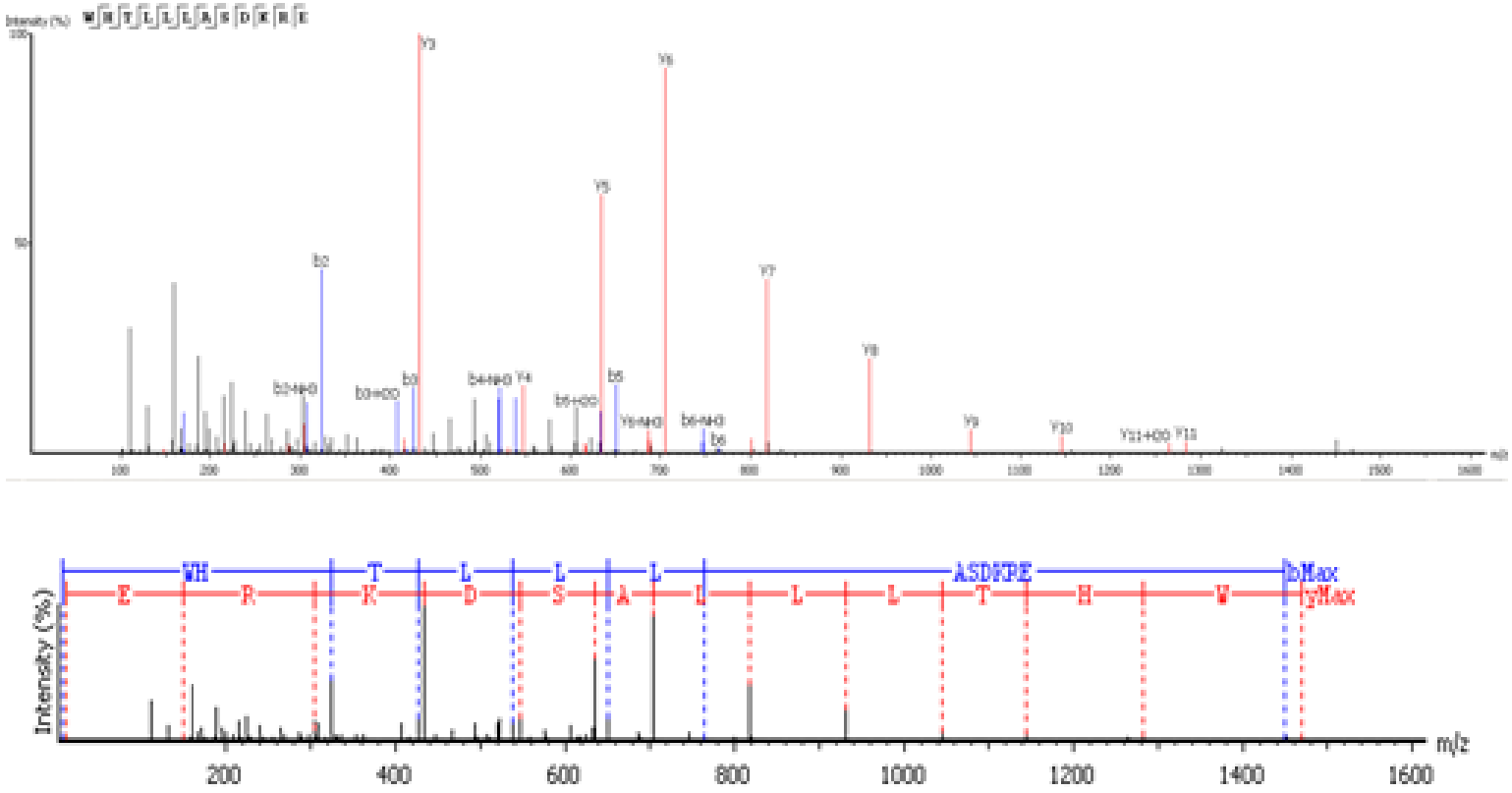
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1439 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



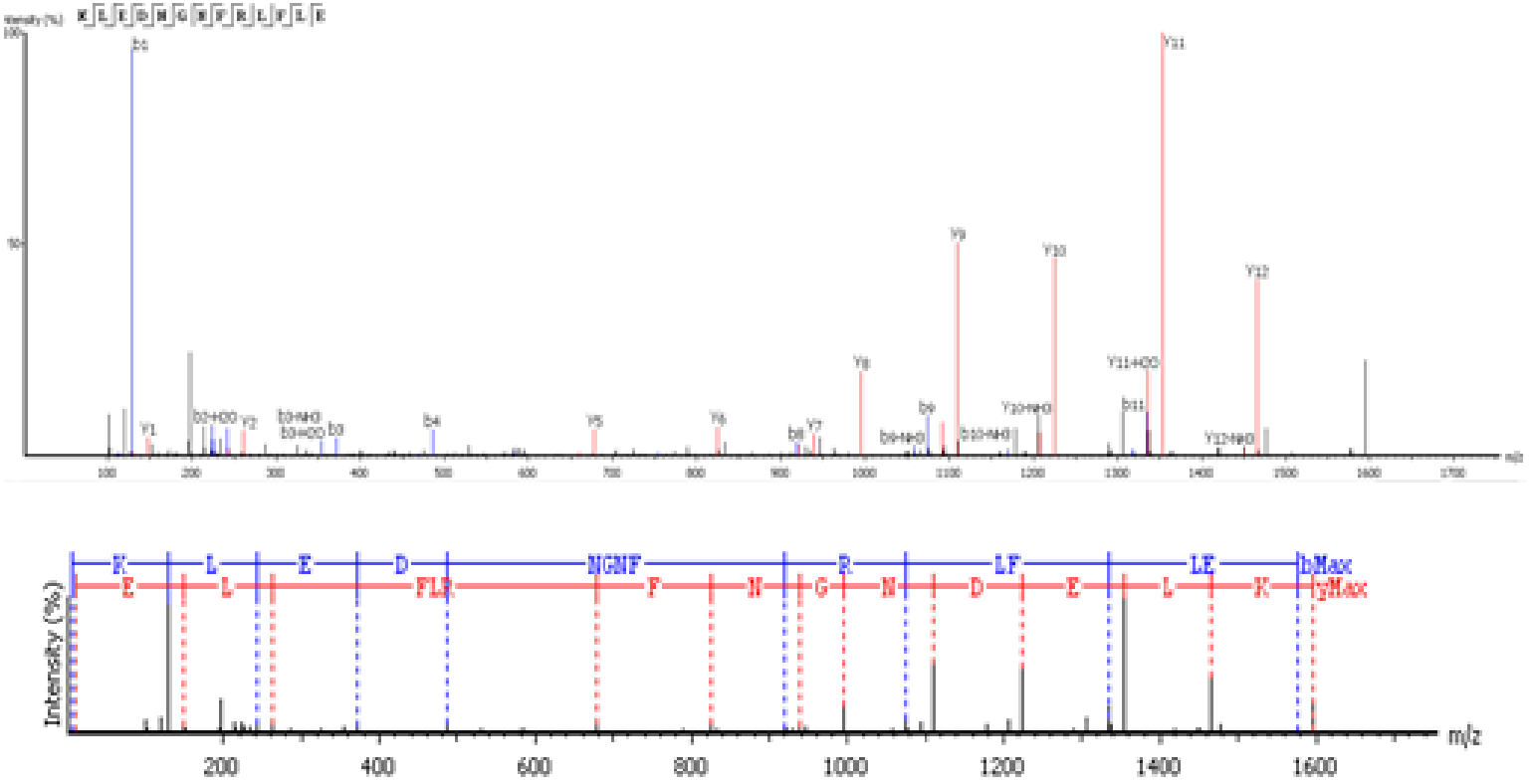
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1049 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



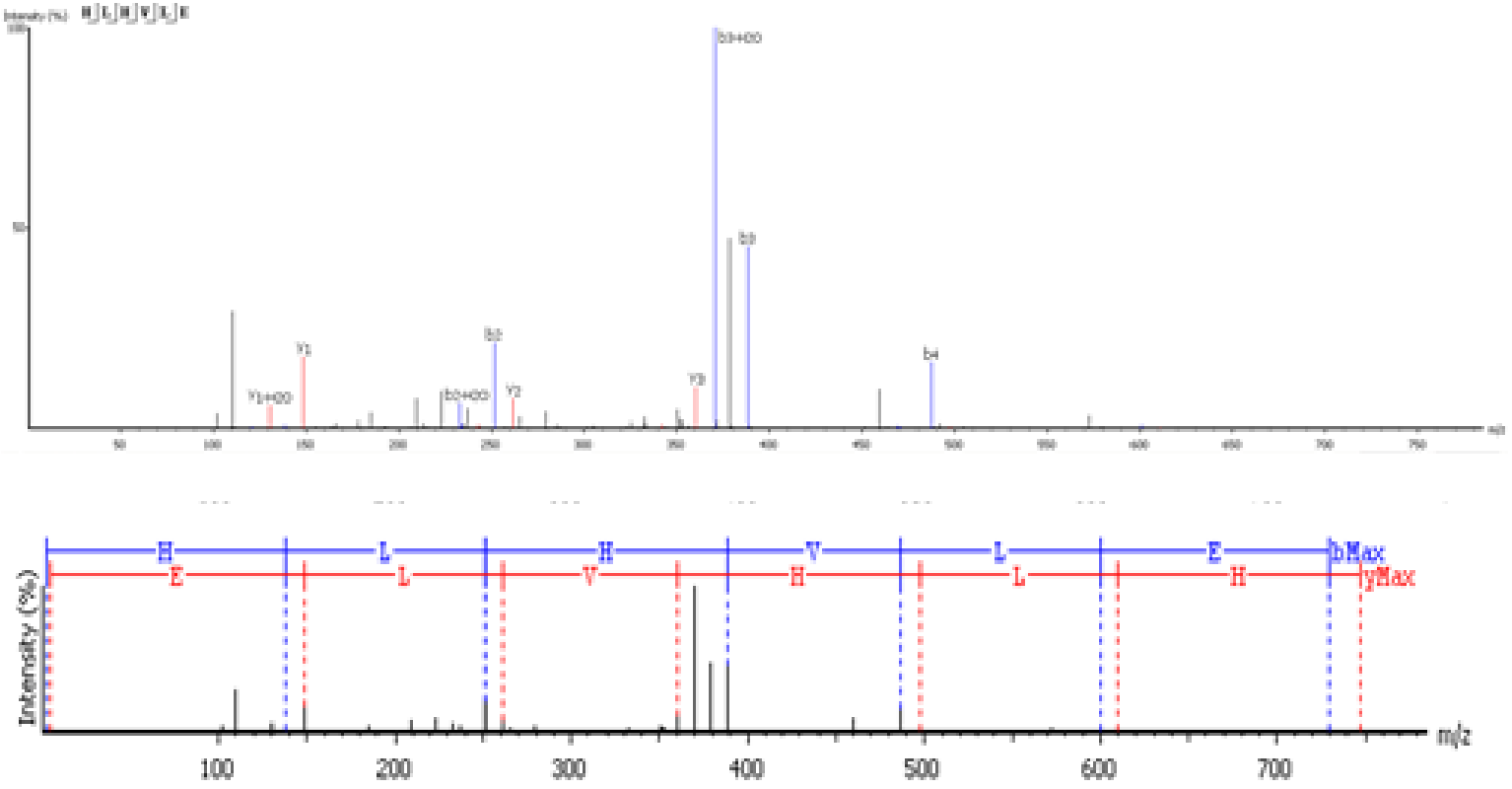
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1468 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



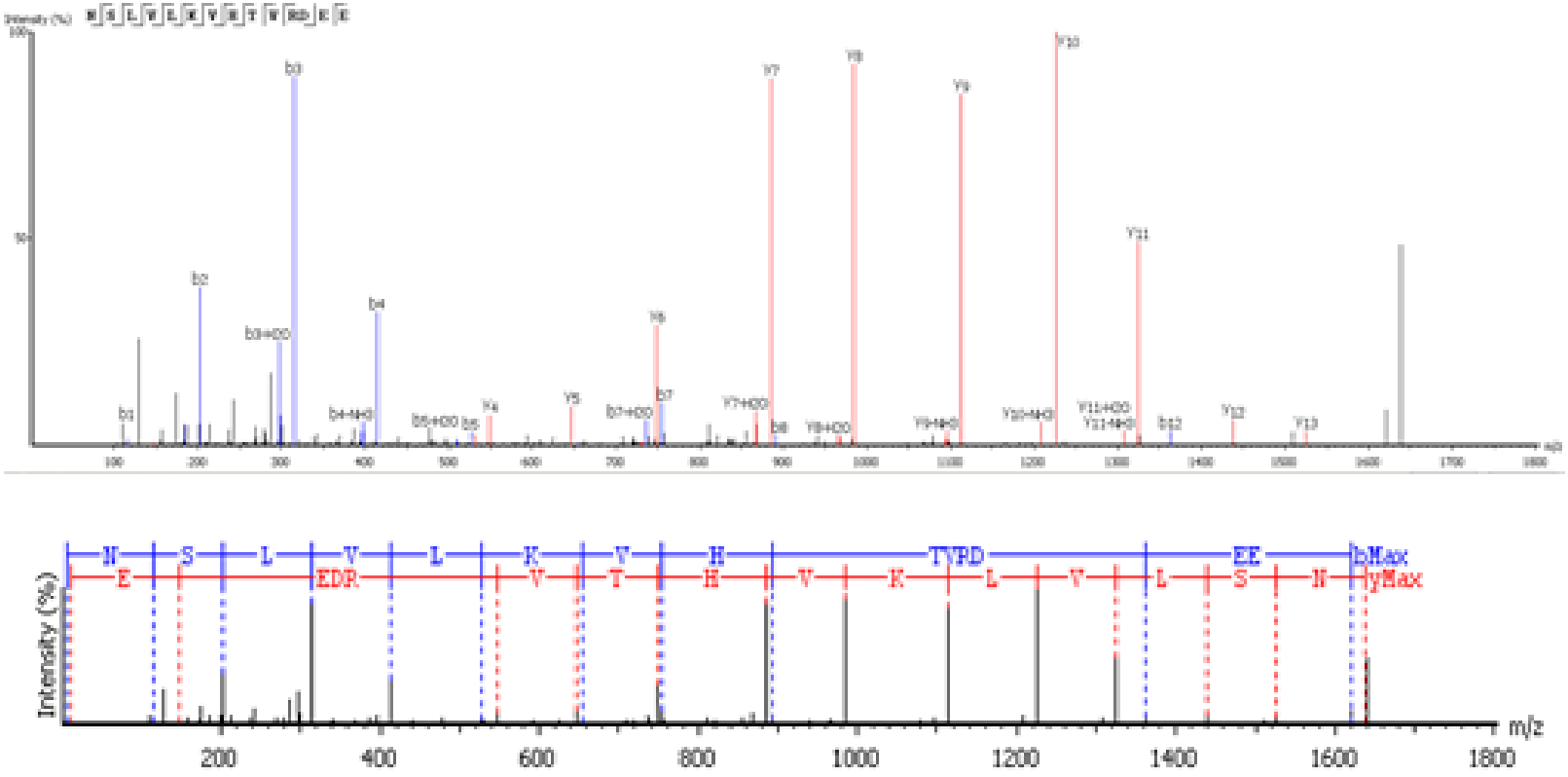
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1594 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



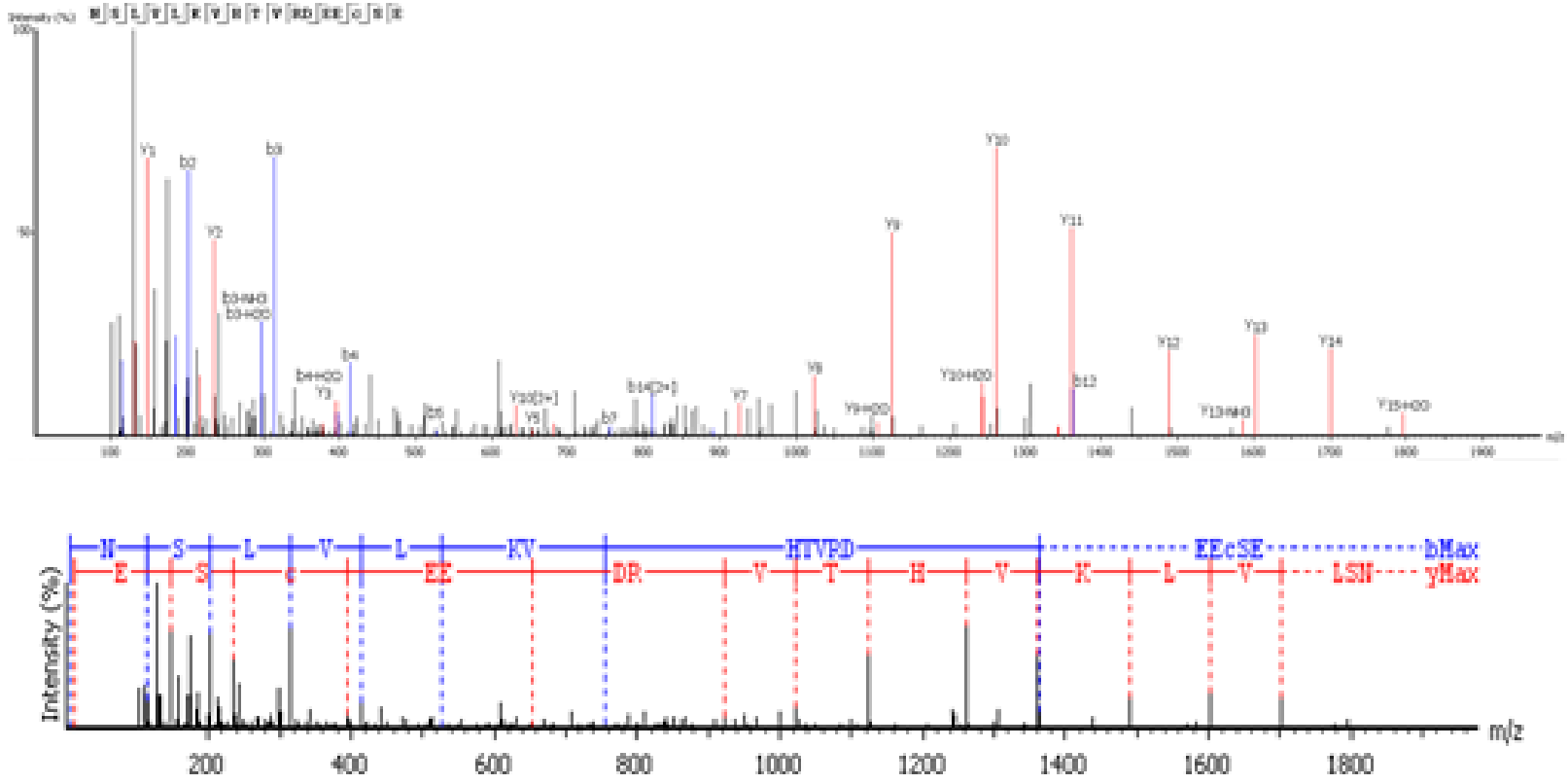
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 746 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



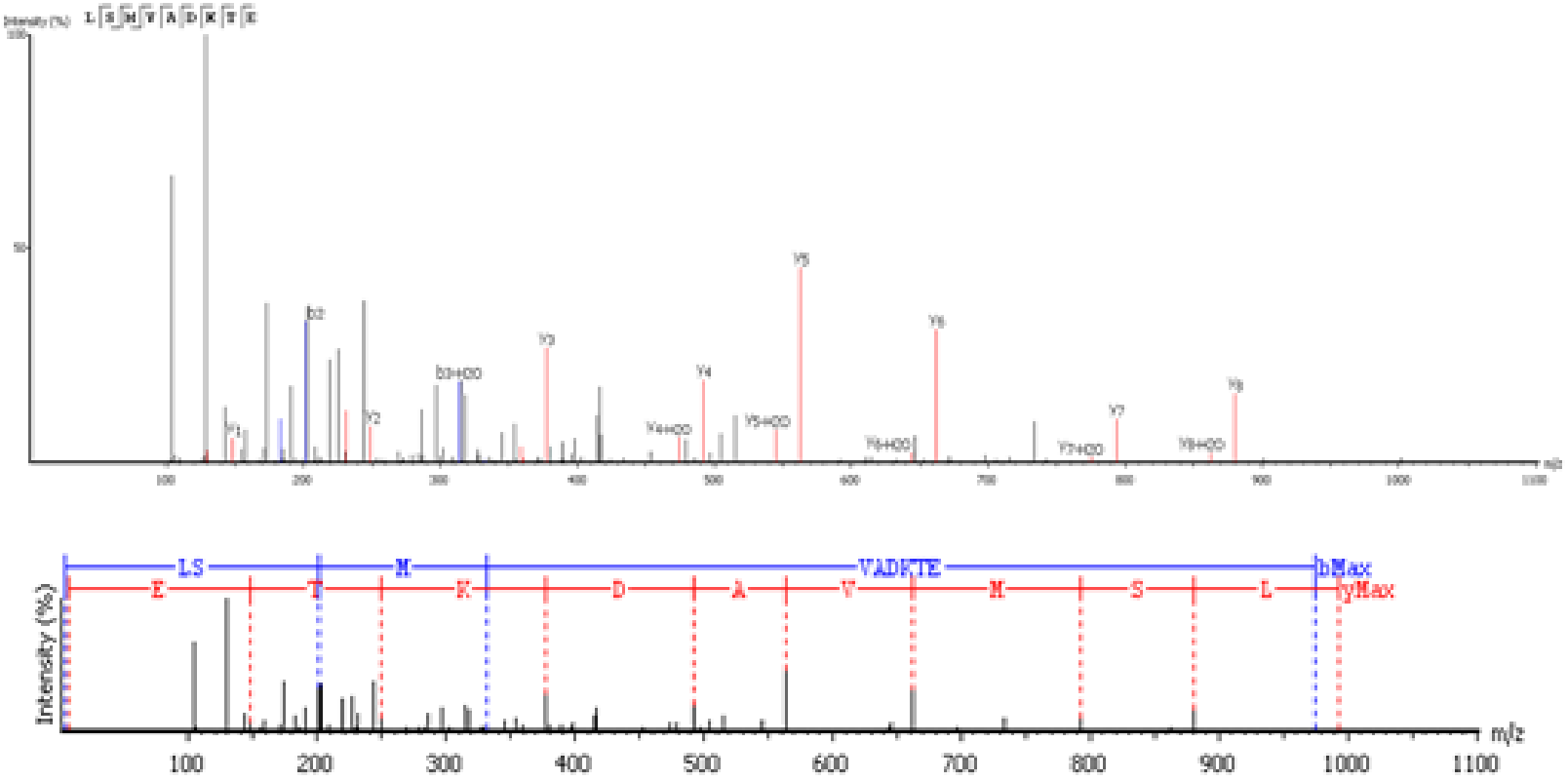
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1638 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



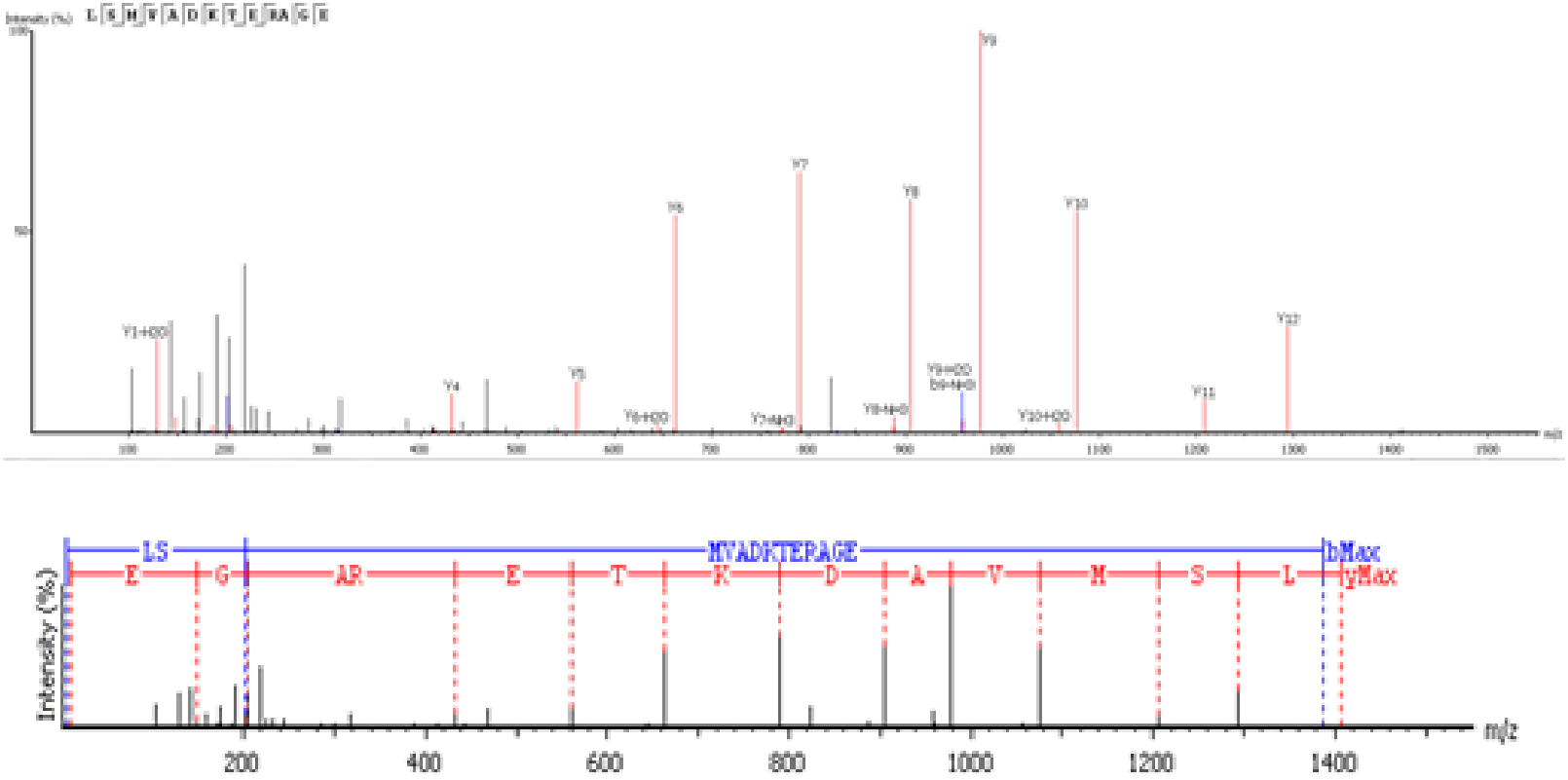
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 2014 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



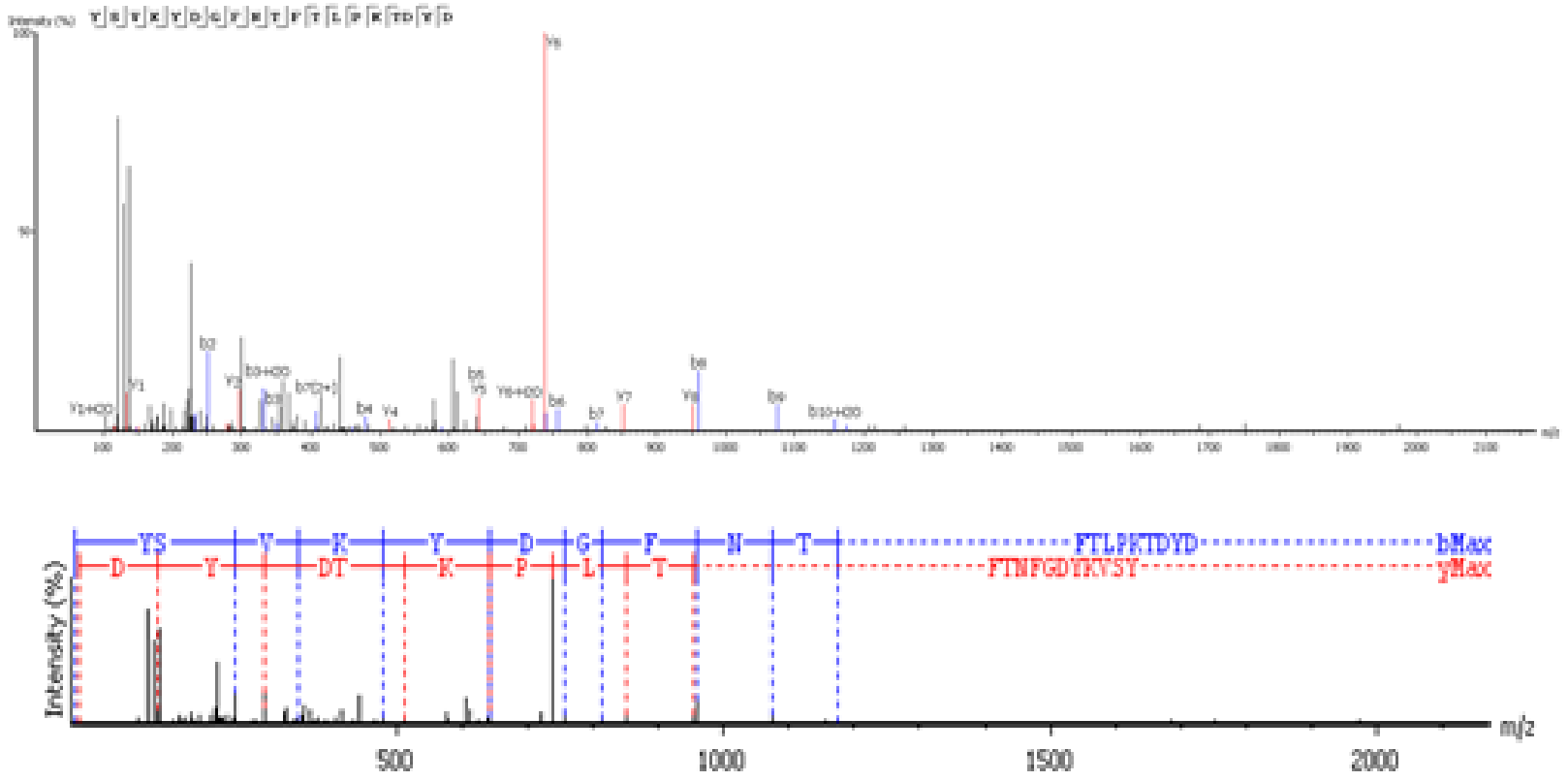
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 992 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



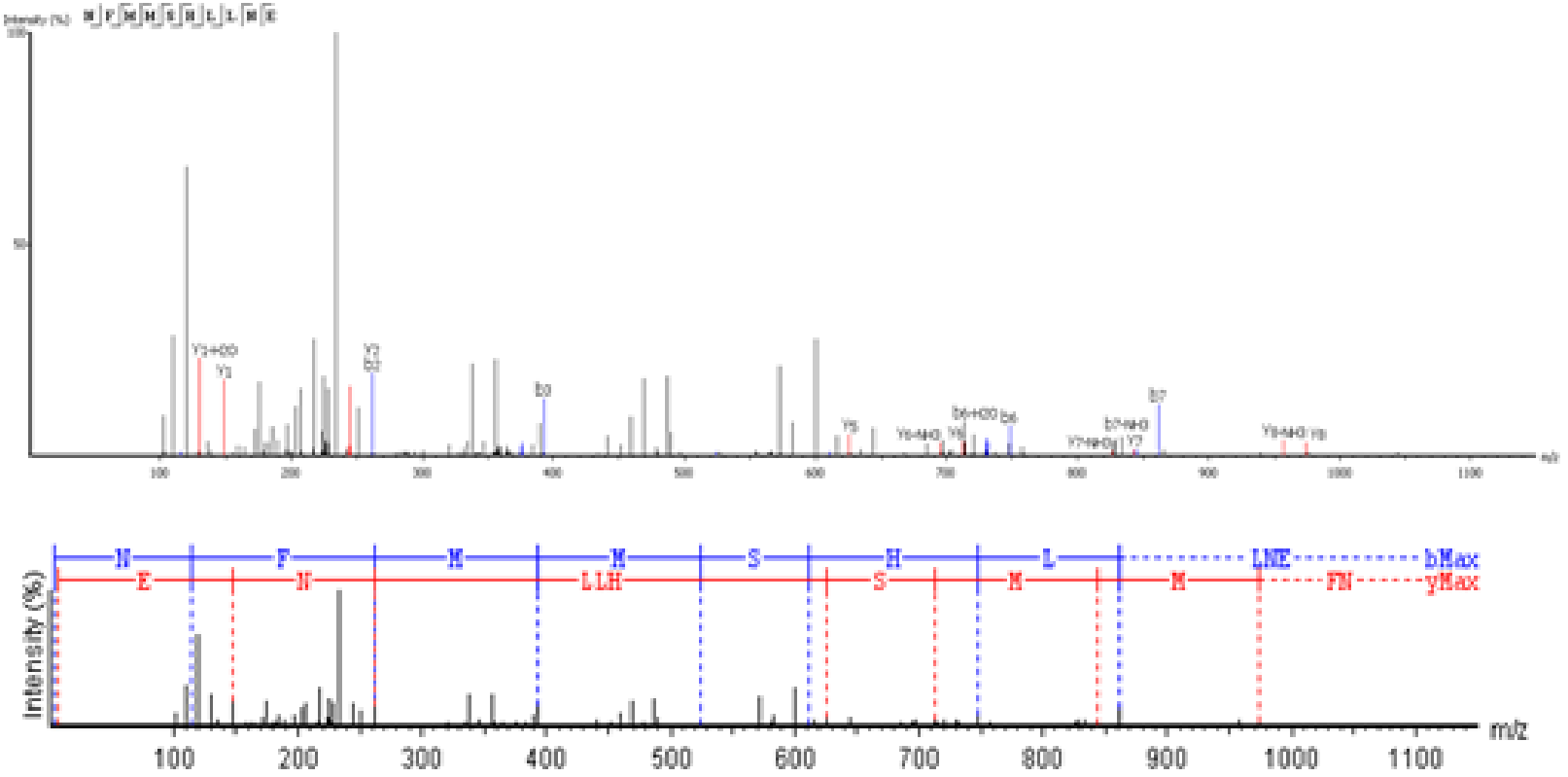
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1406 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



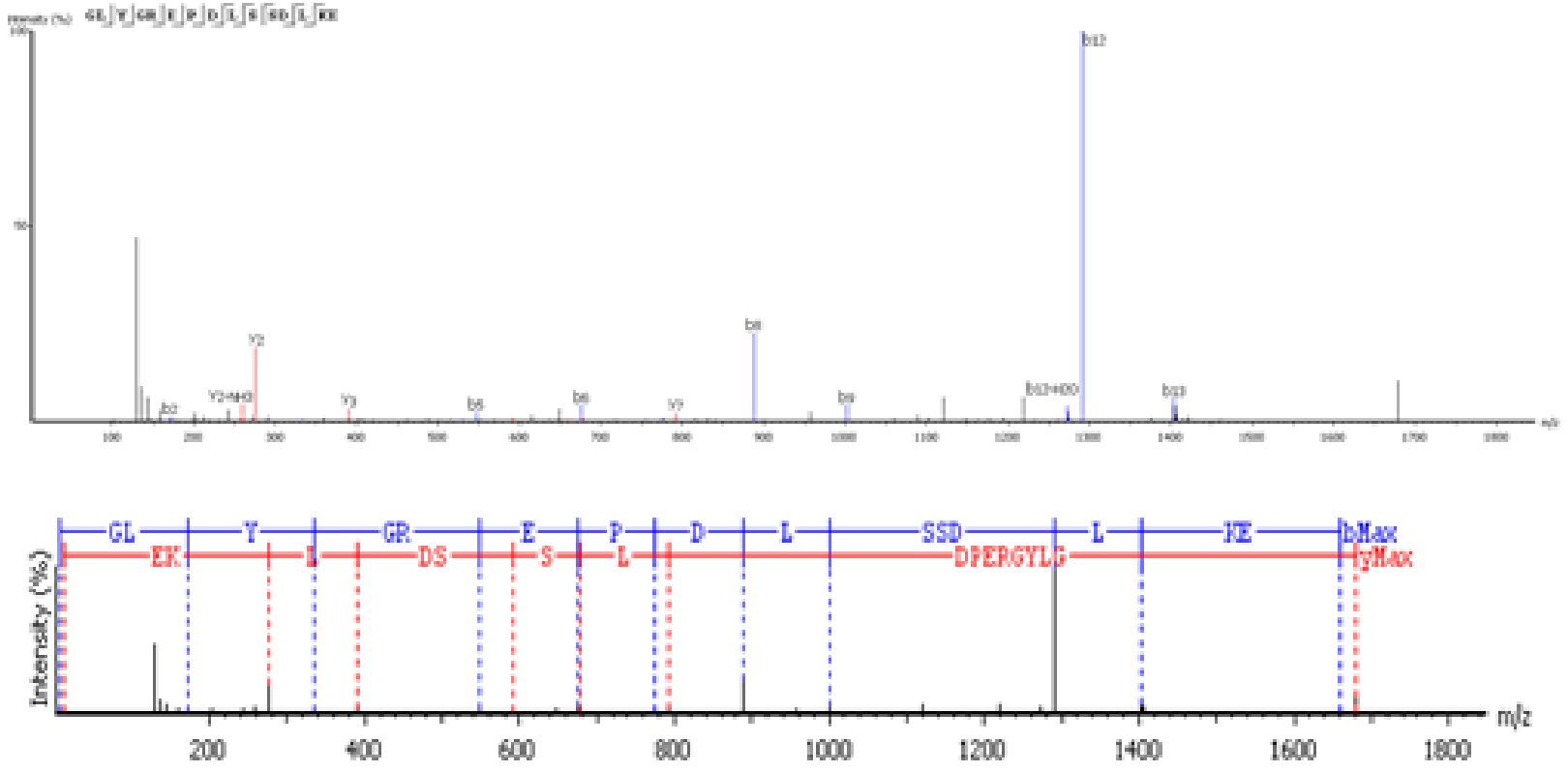
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 2273 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



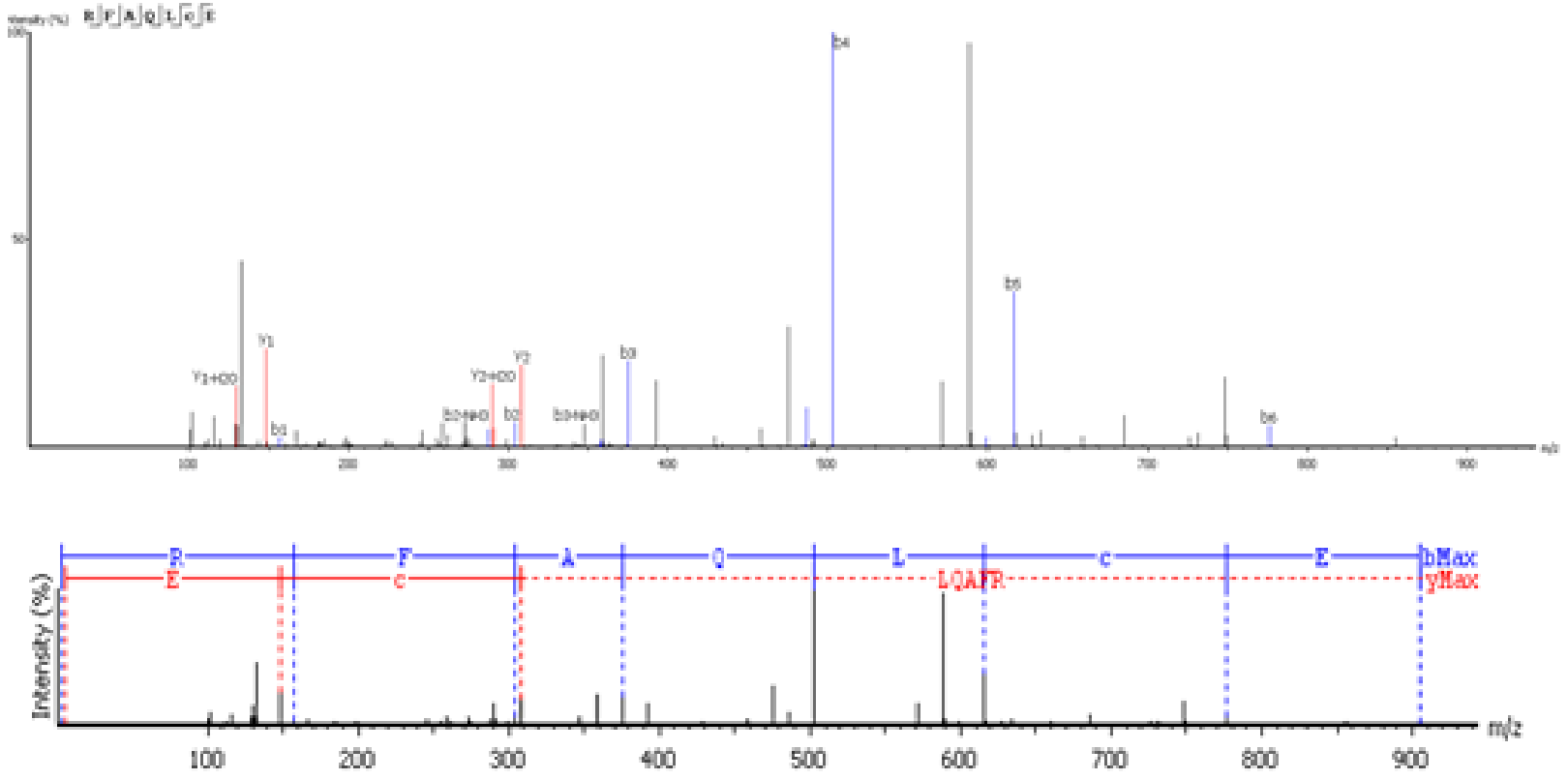
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1235 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



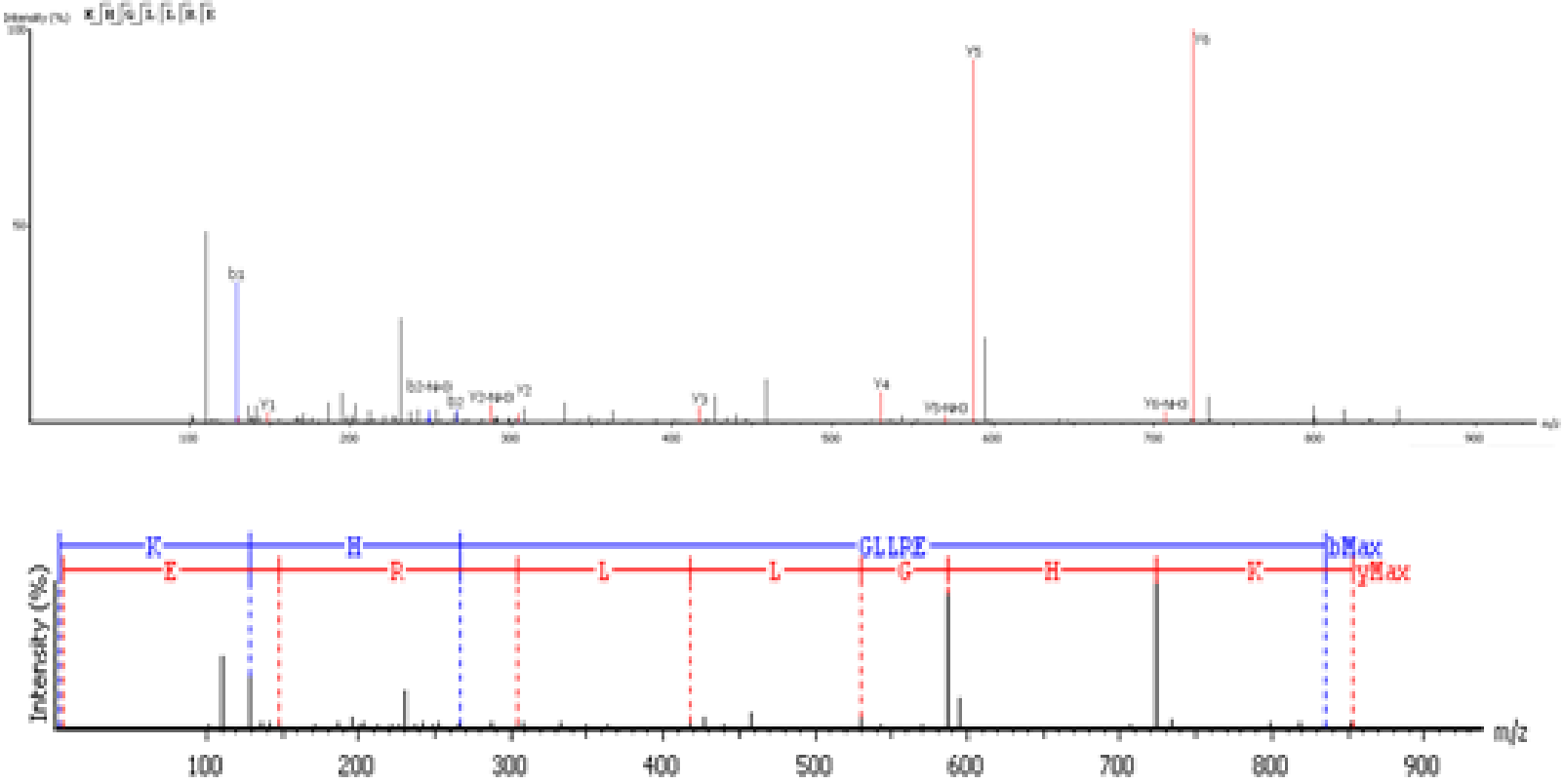
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1678 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



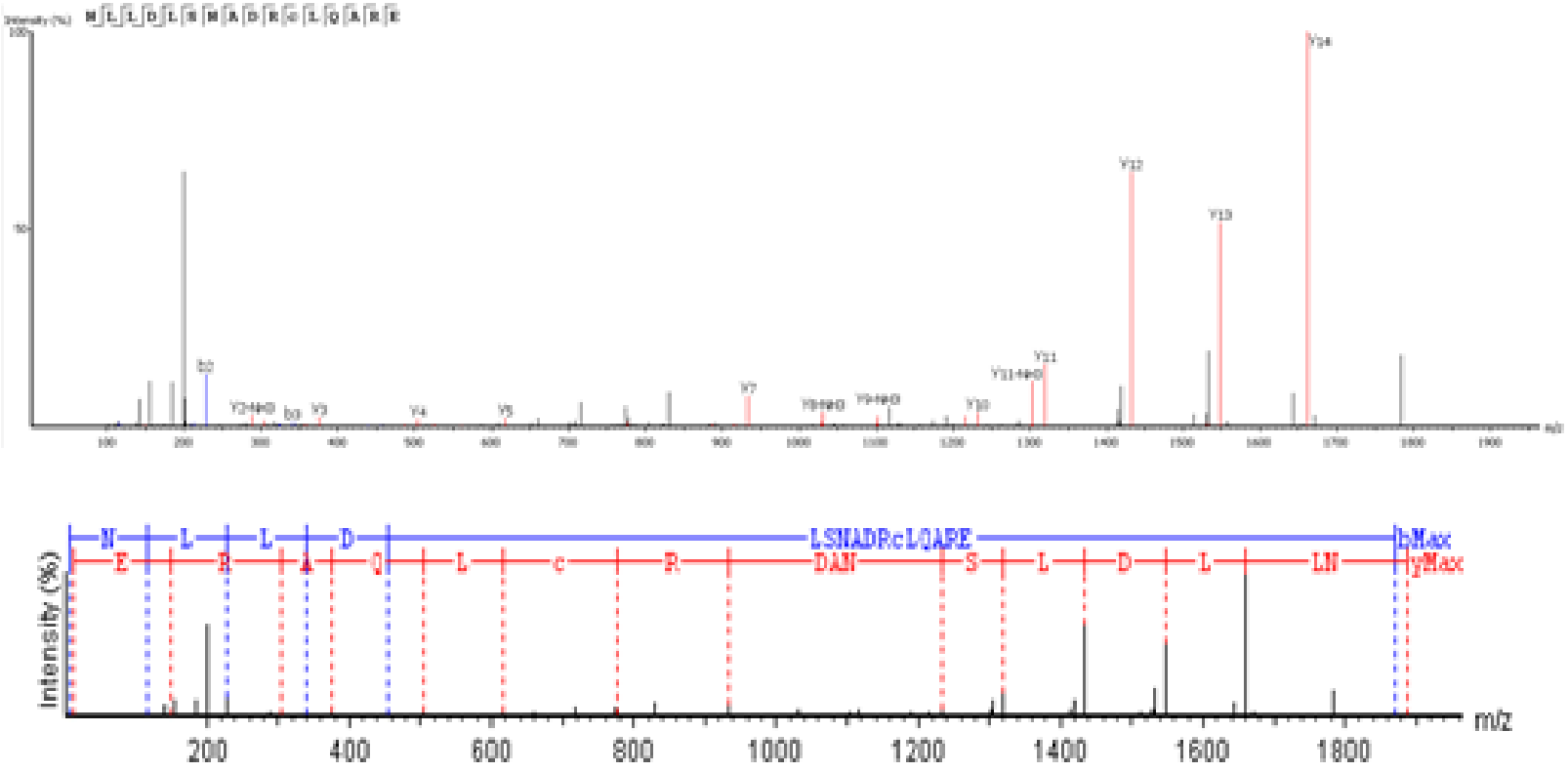
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 922 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



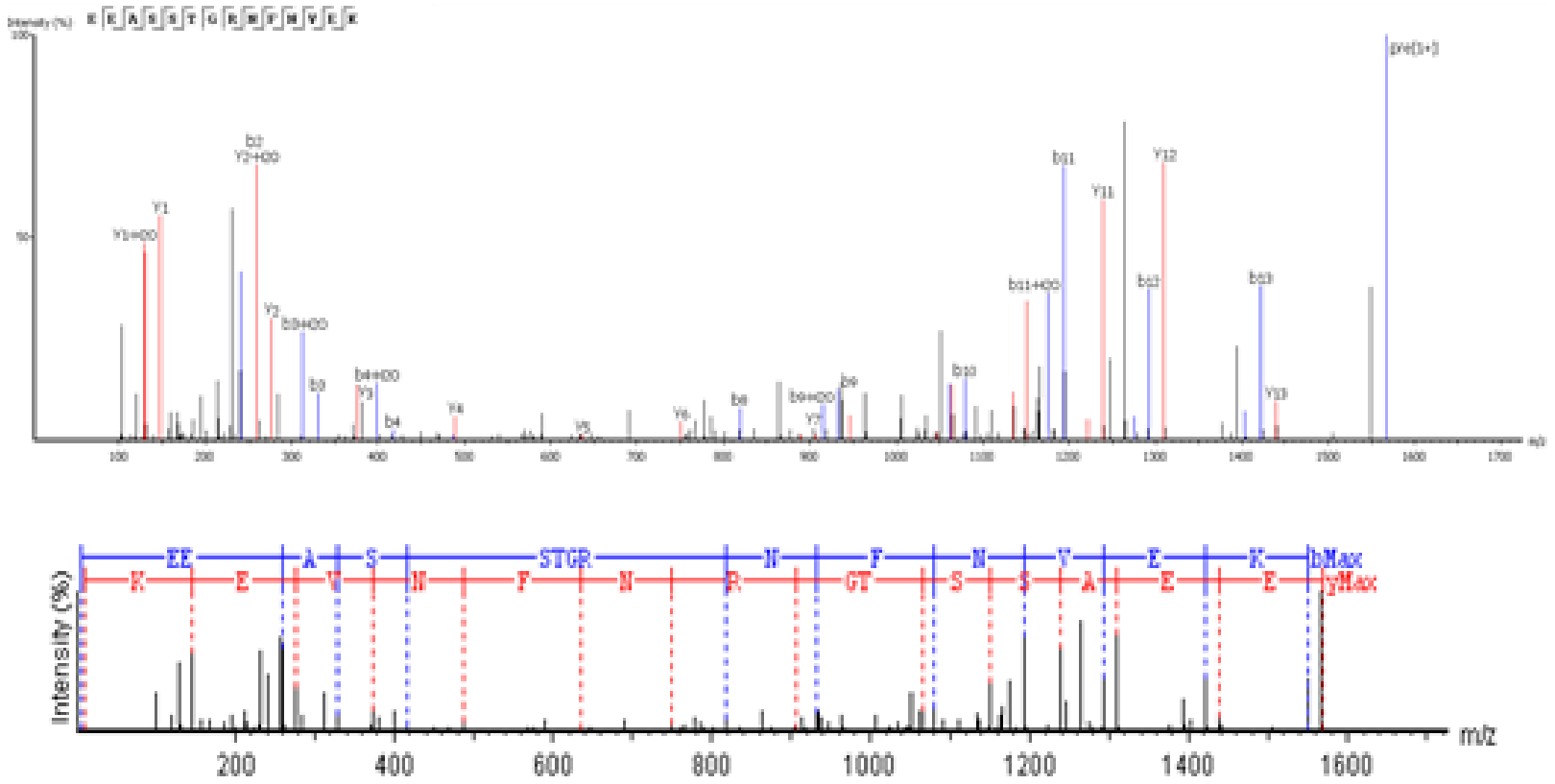
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 851 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



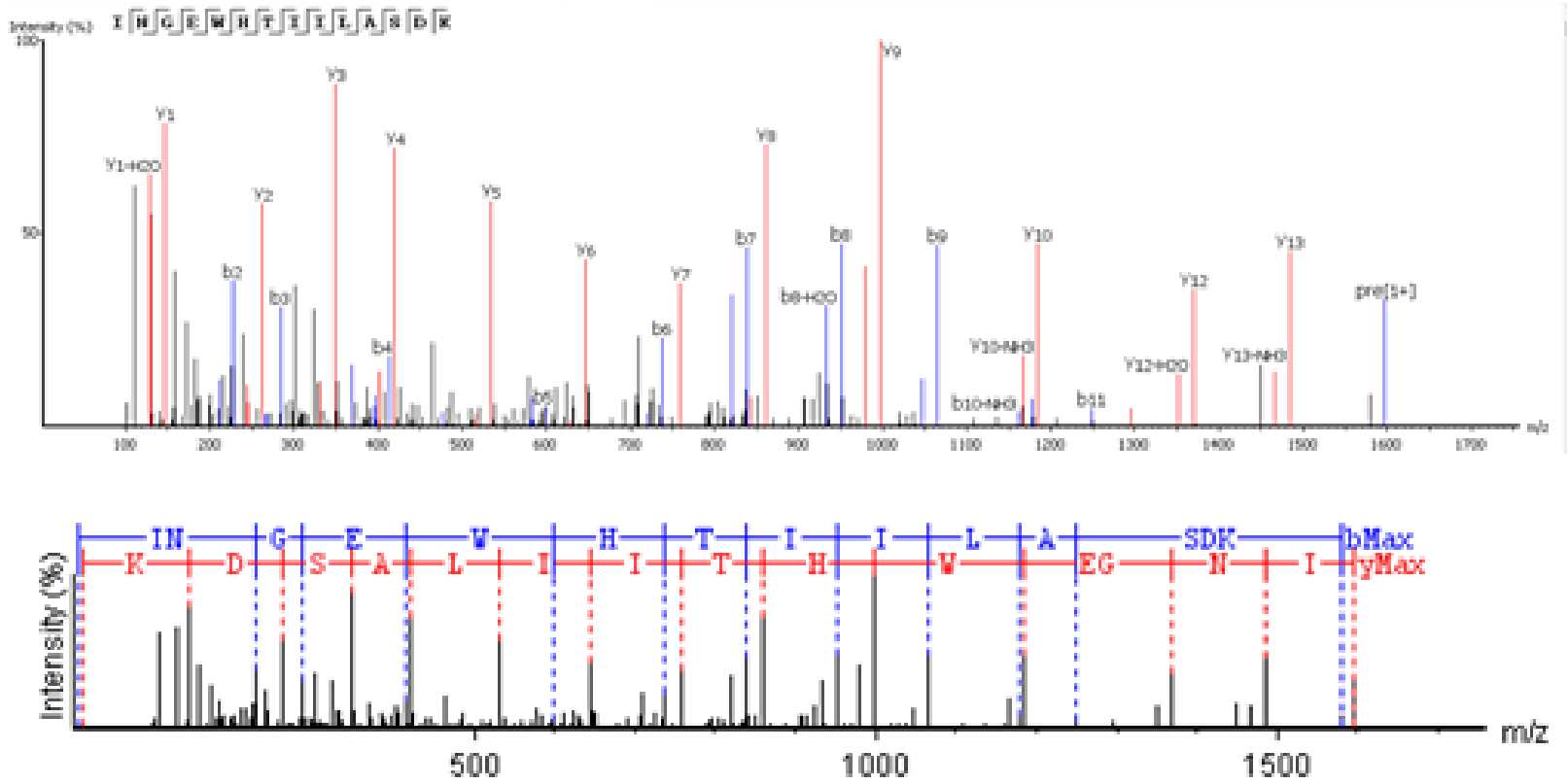
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1887 Da of 18742 Da MUP.

Fraction 1 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



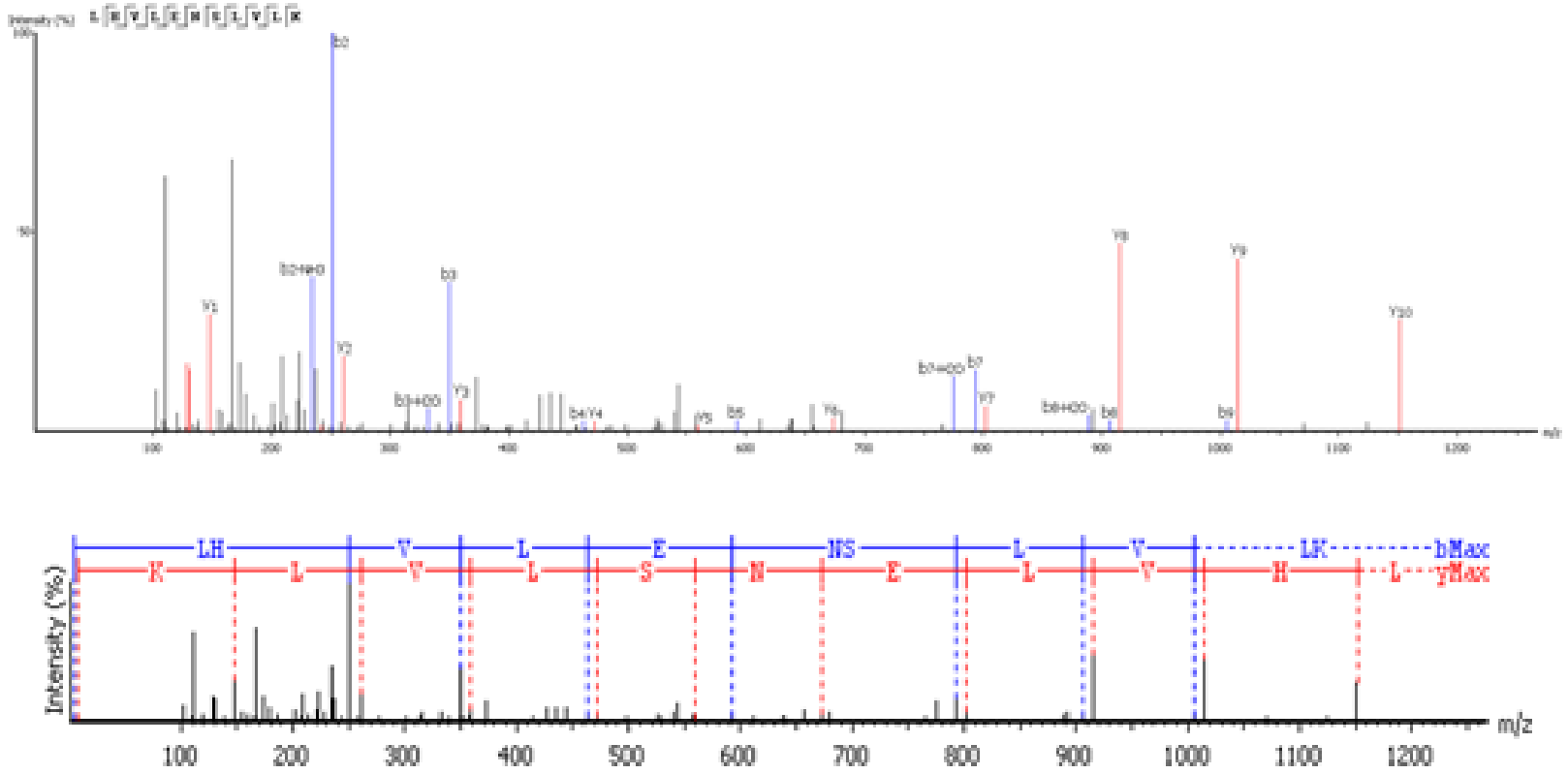
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1567 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



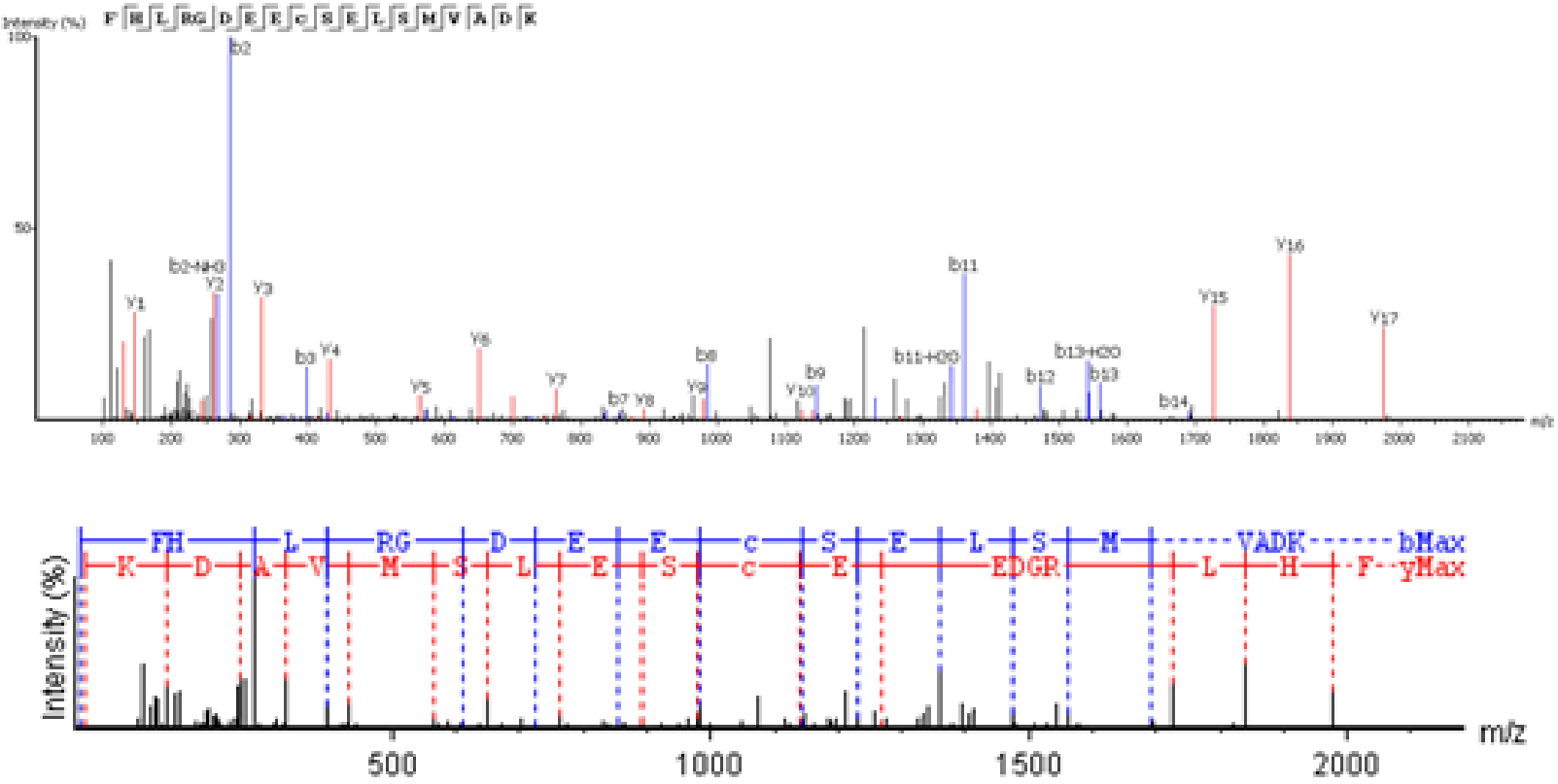
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1596 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



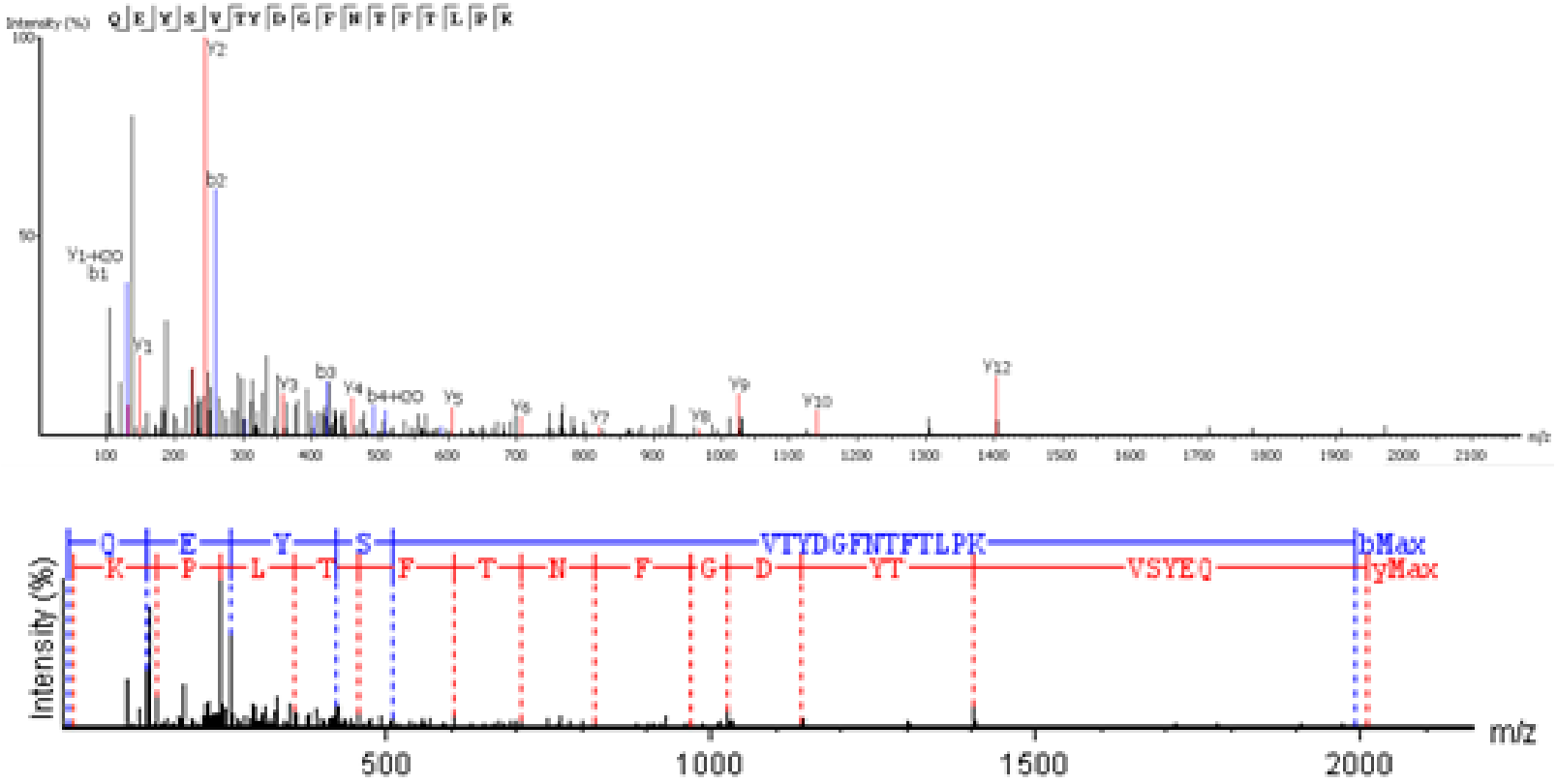
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1264 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



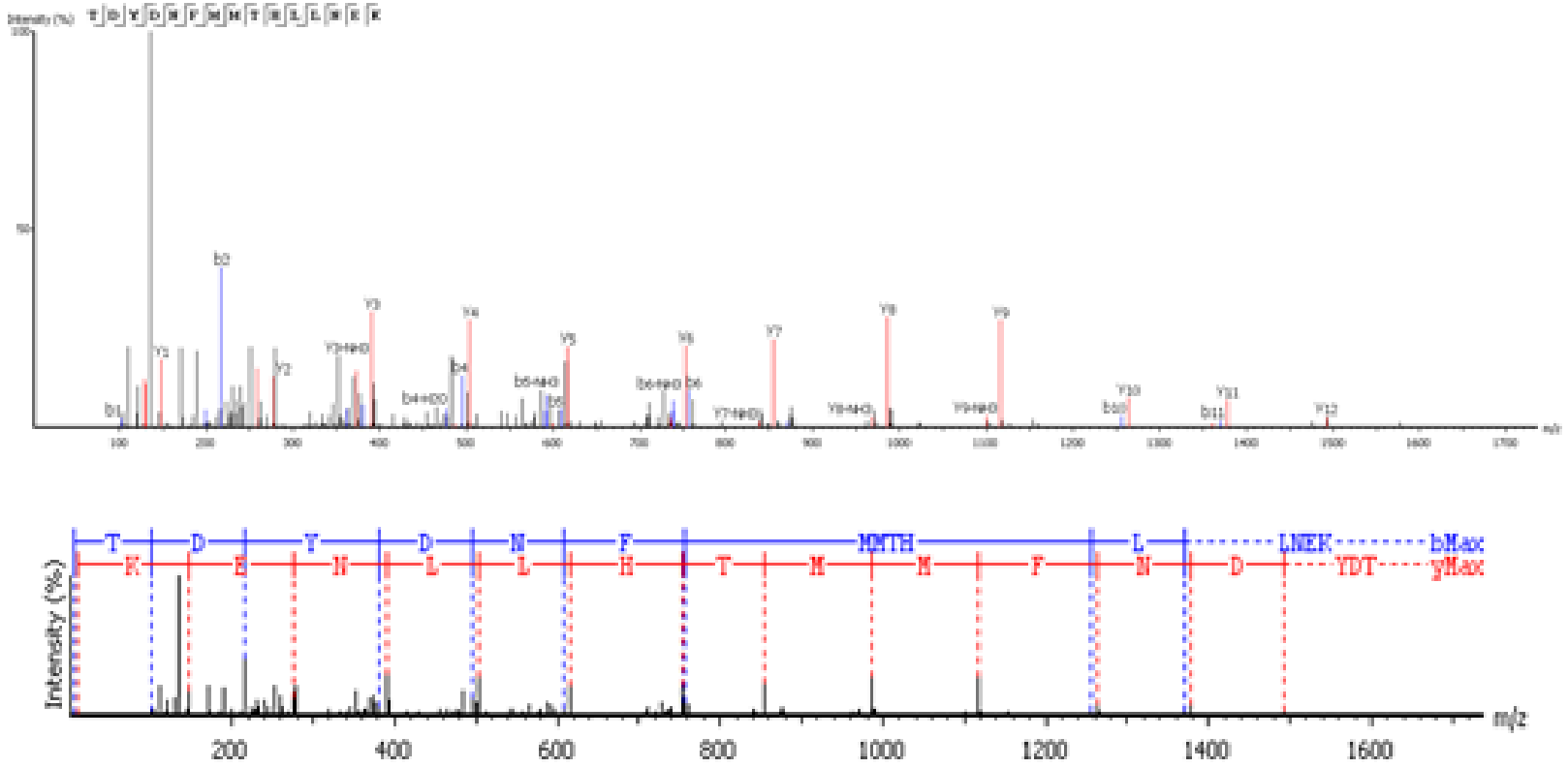
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2122 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



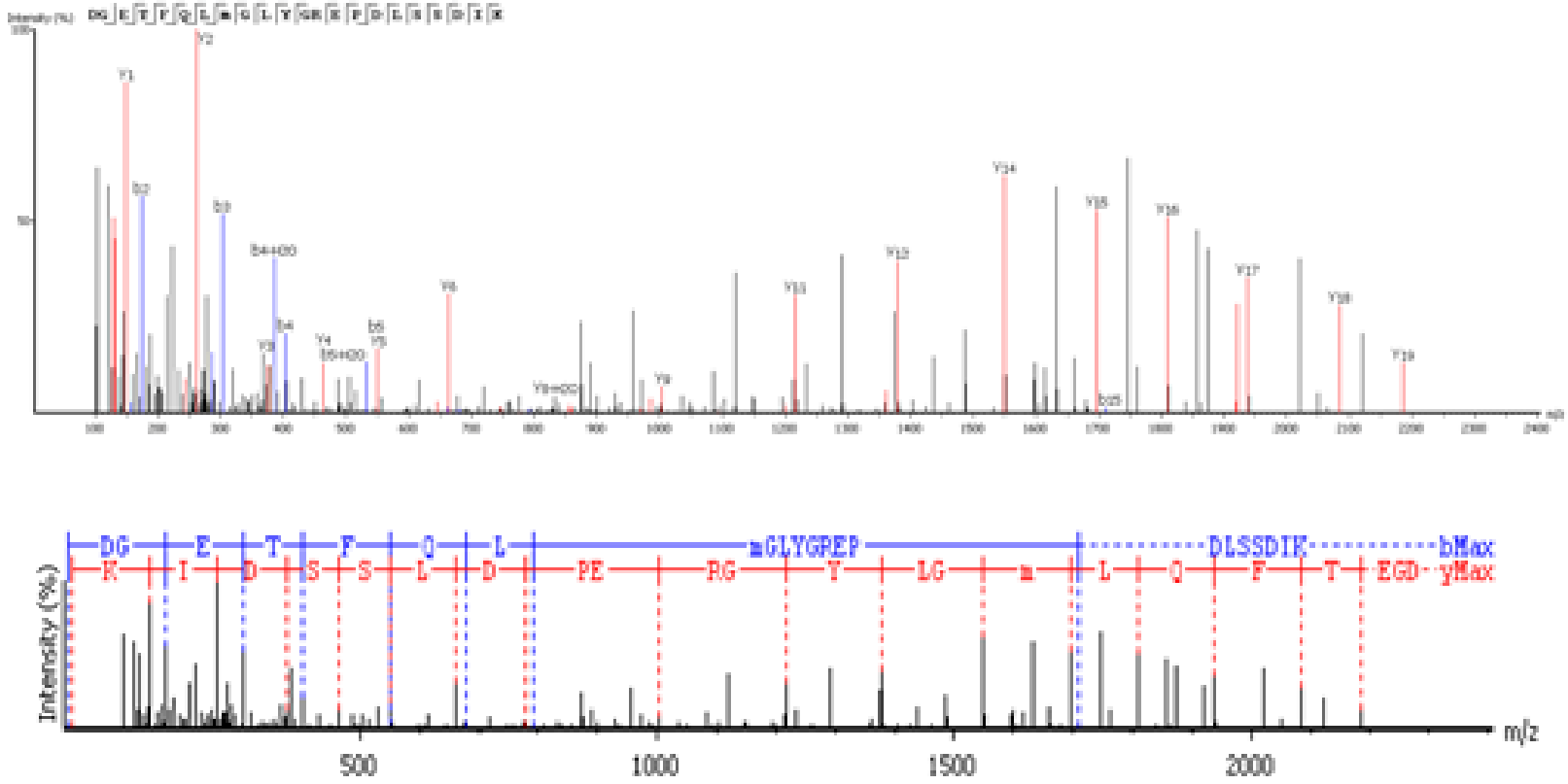
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2009 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



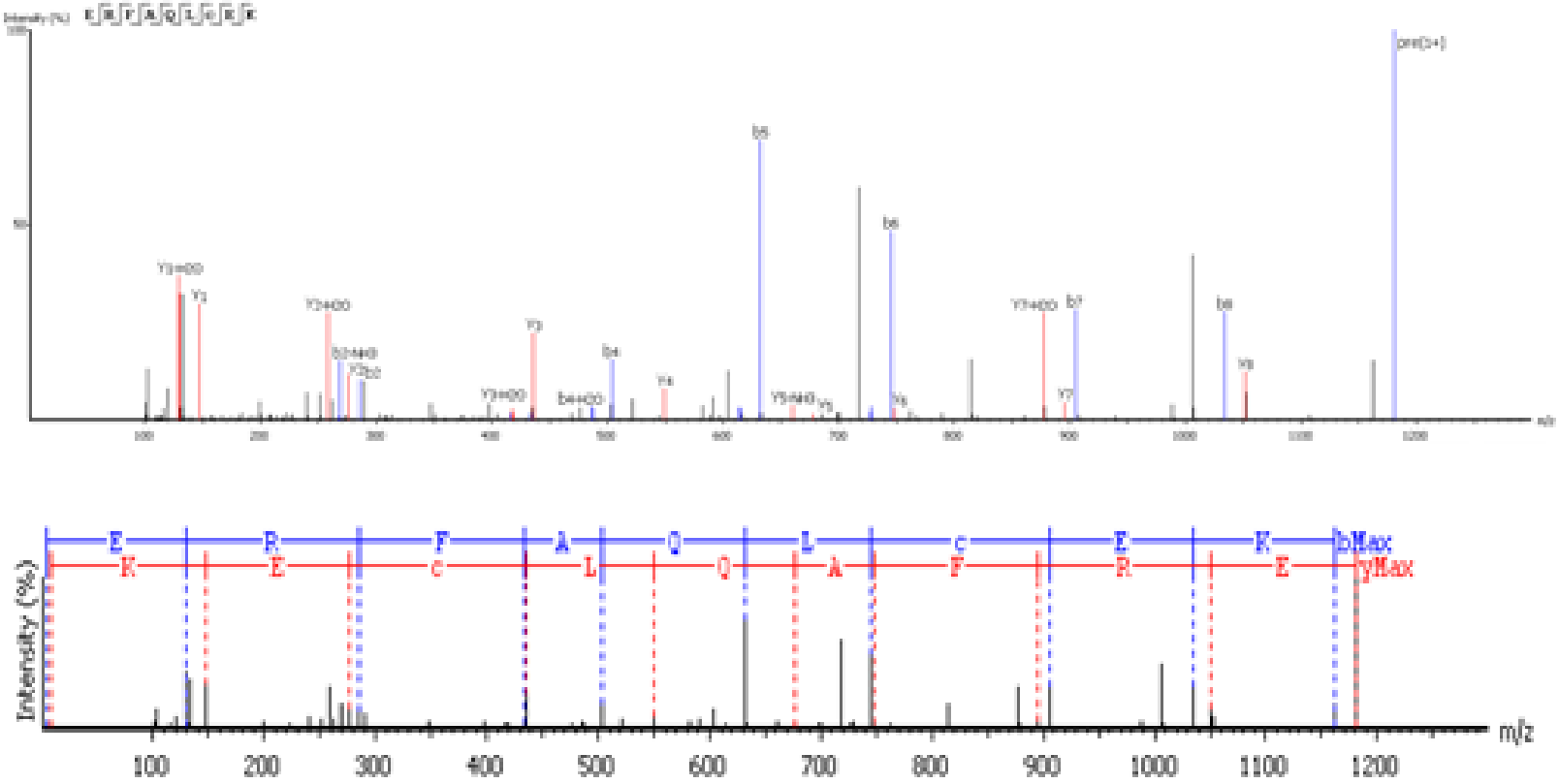
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1871 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



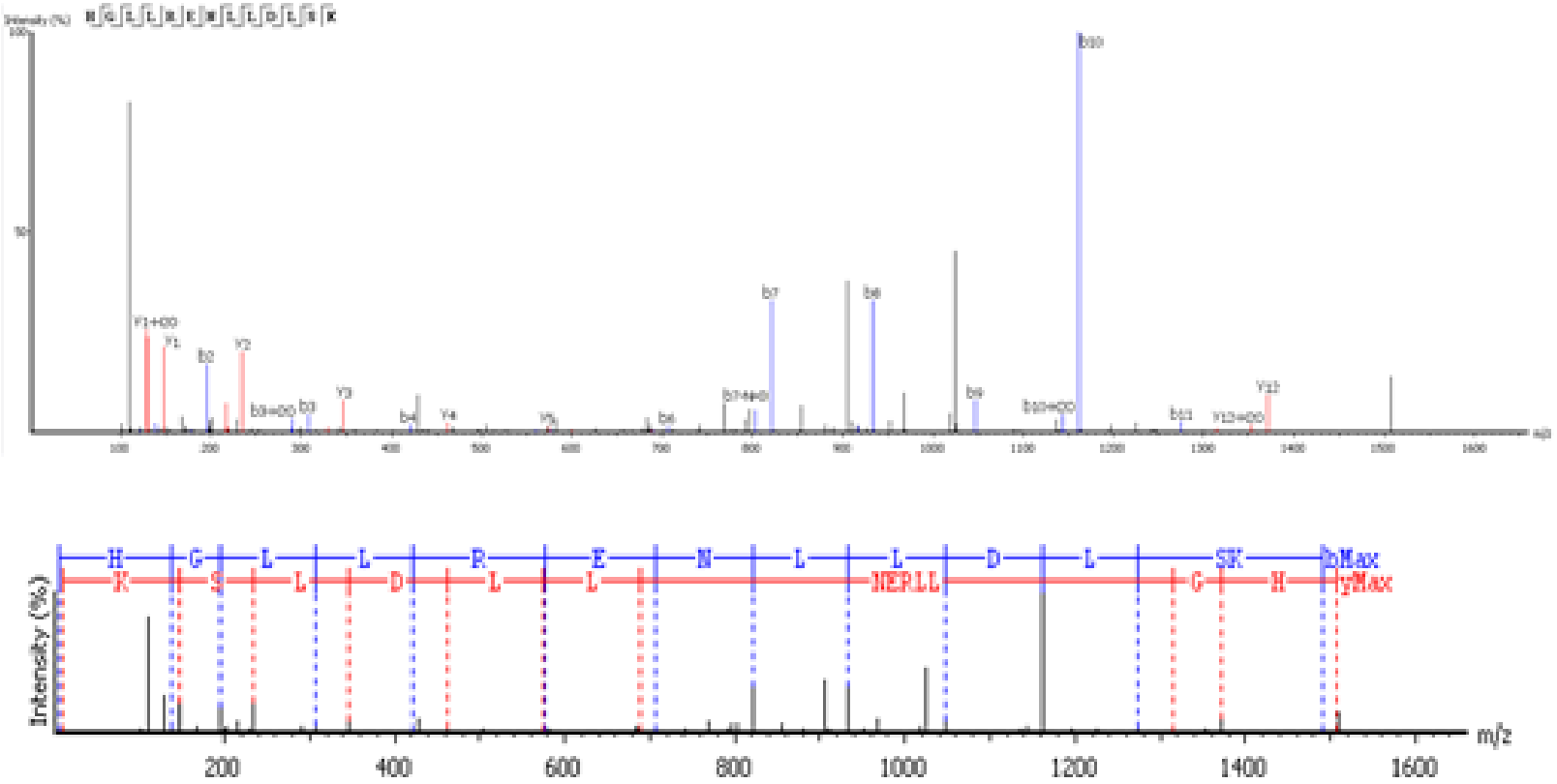
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2486 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



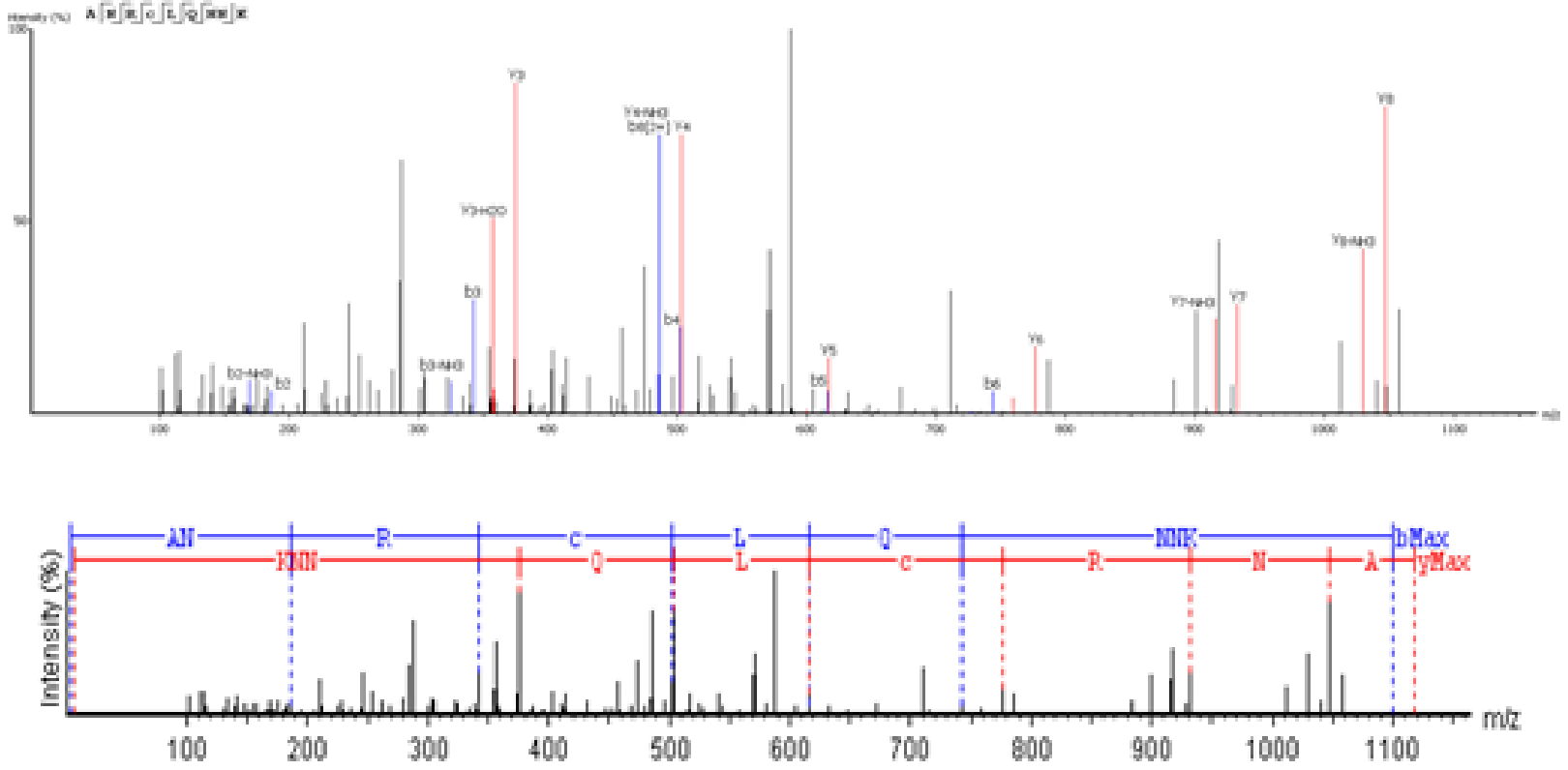
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1180 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



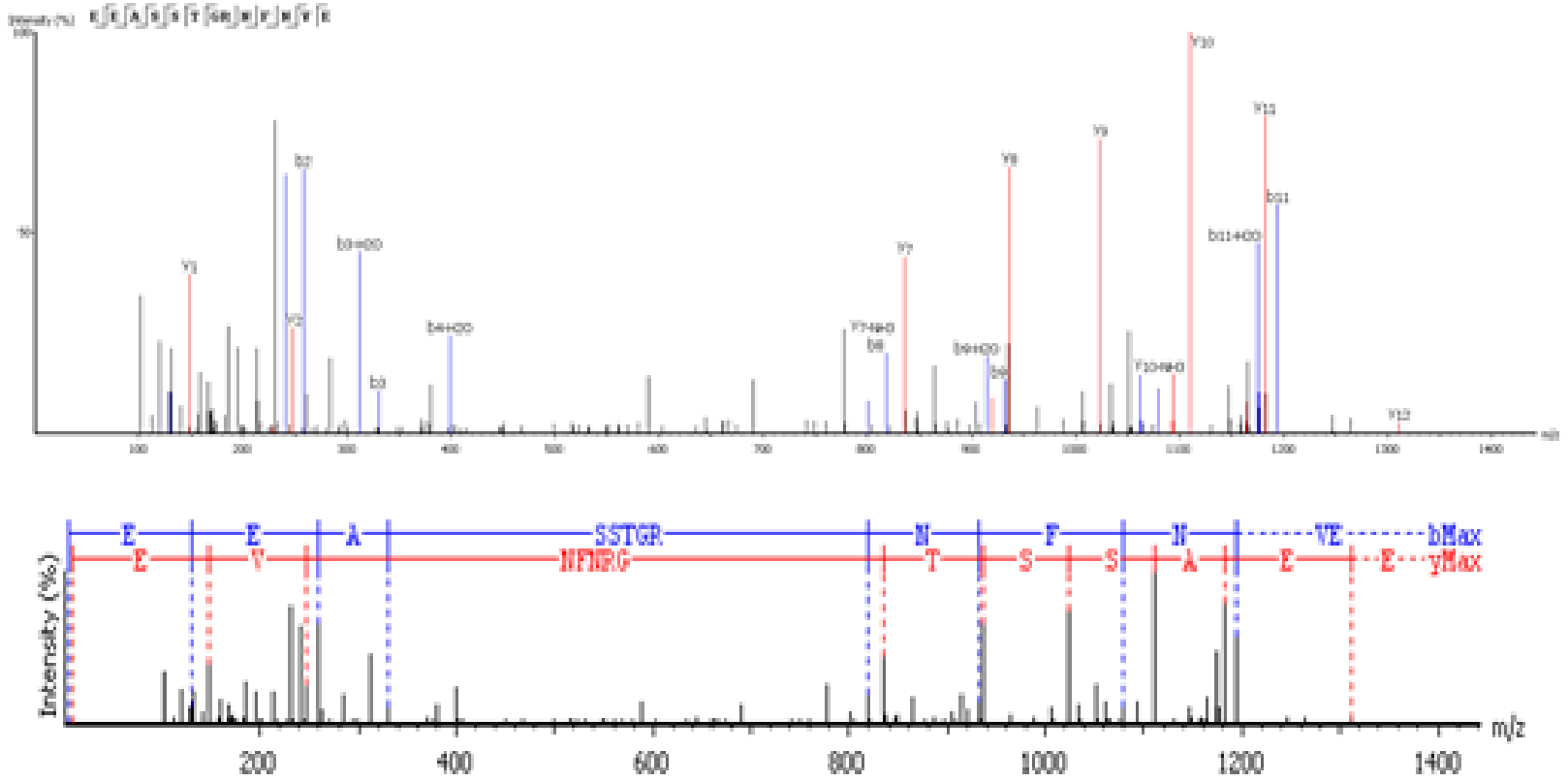
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1507 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



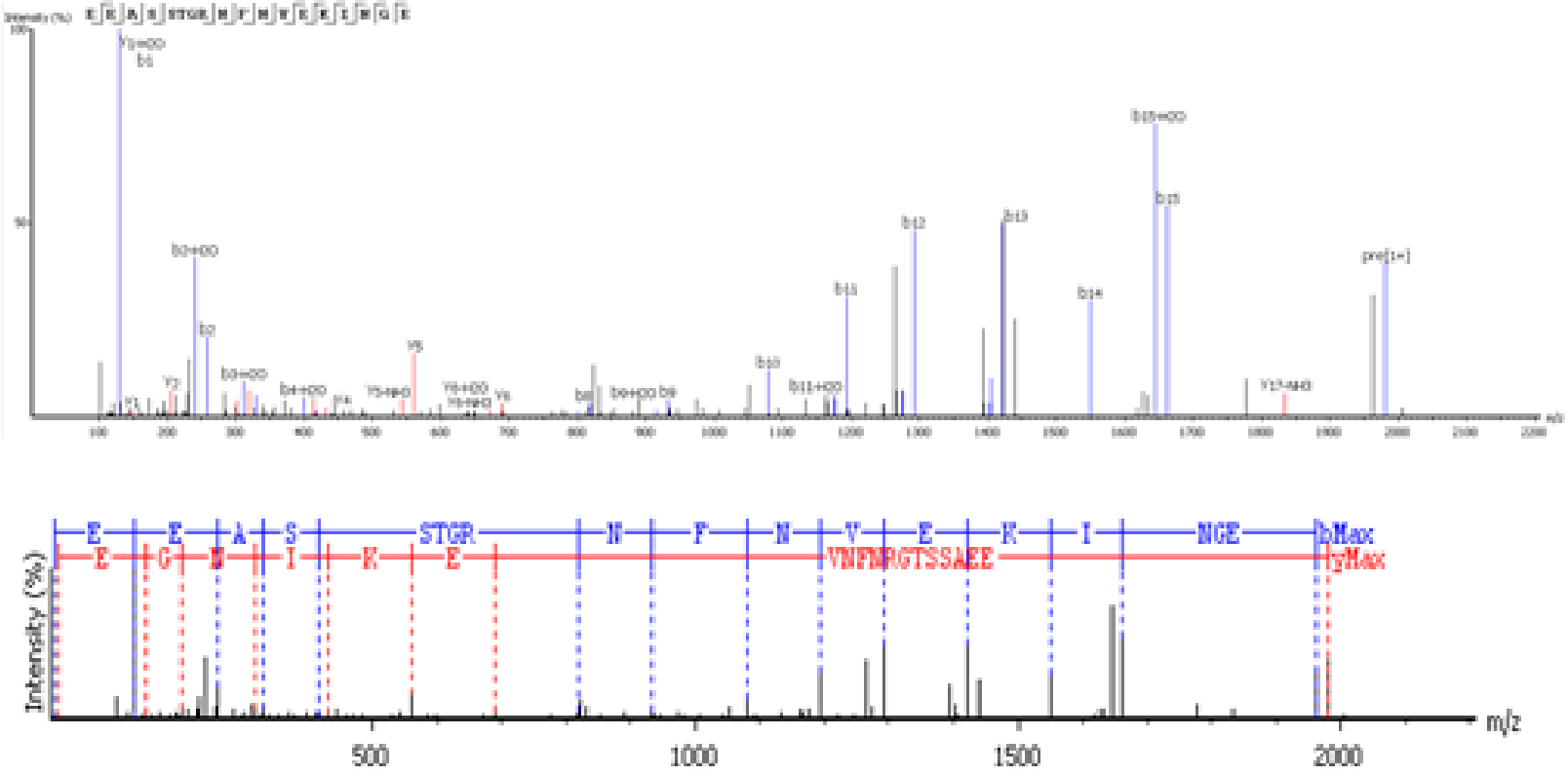
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1117 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



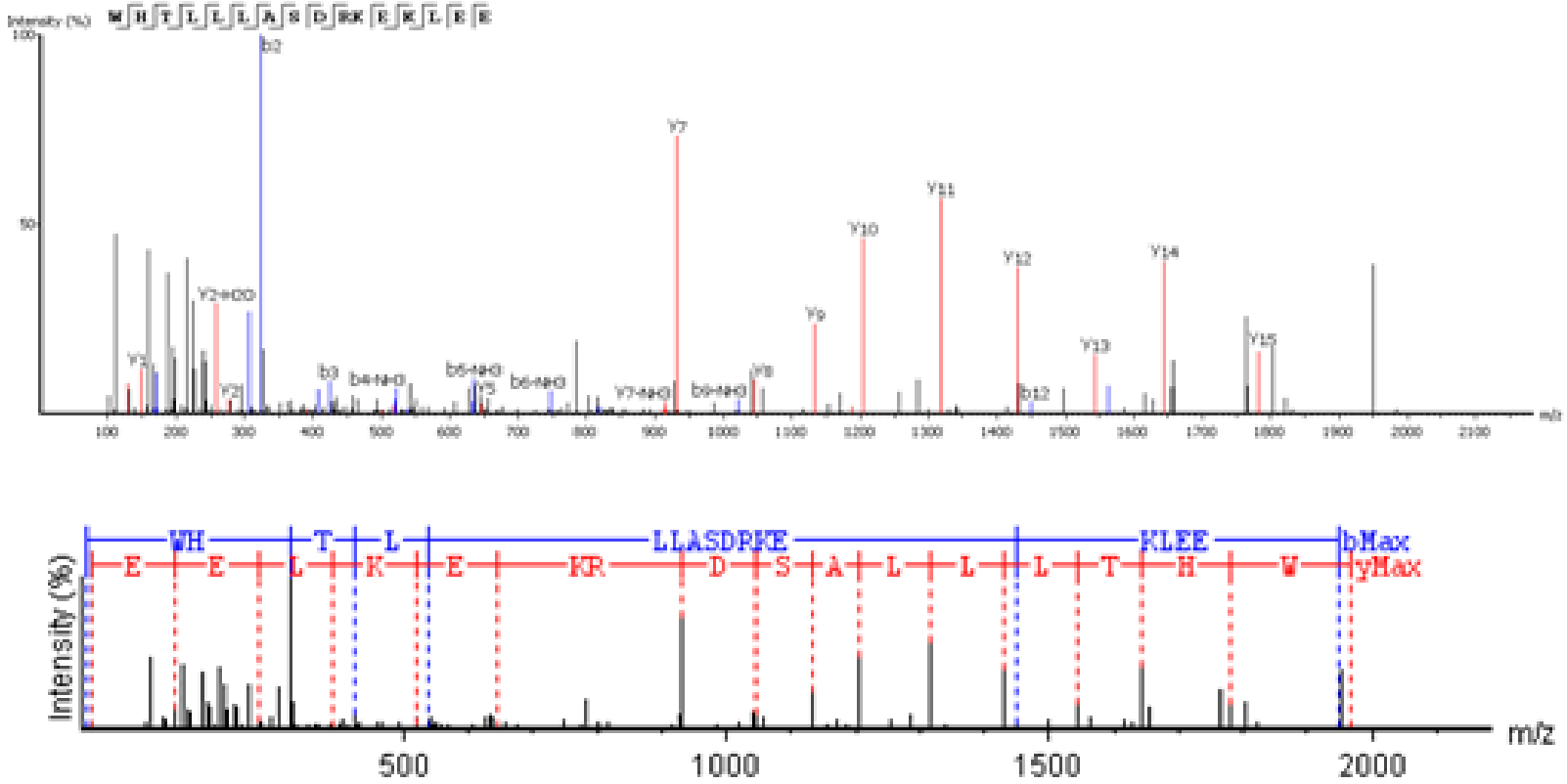
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1439 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



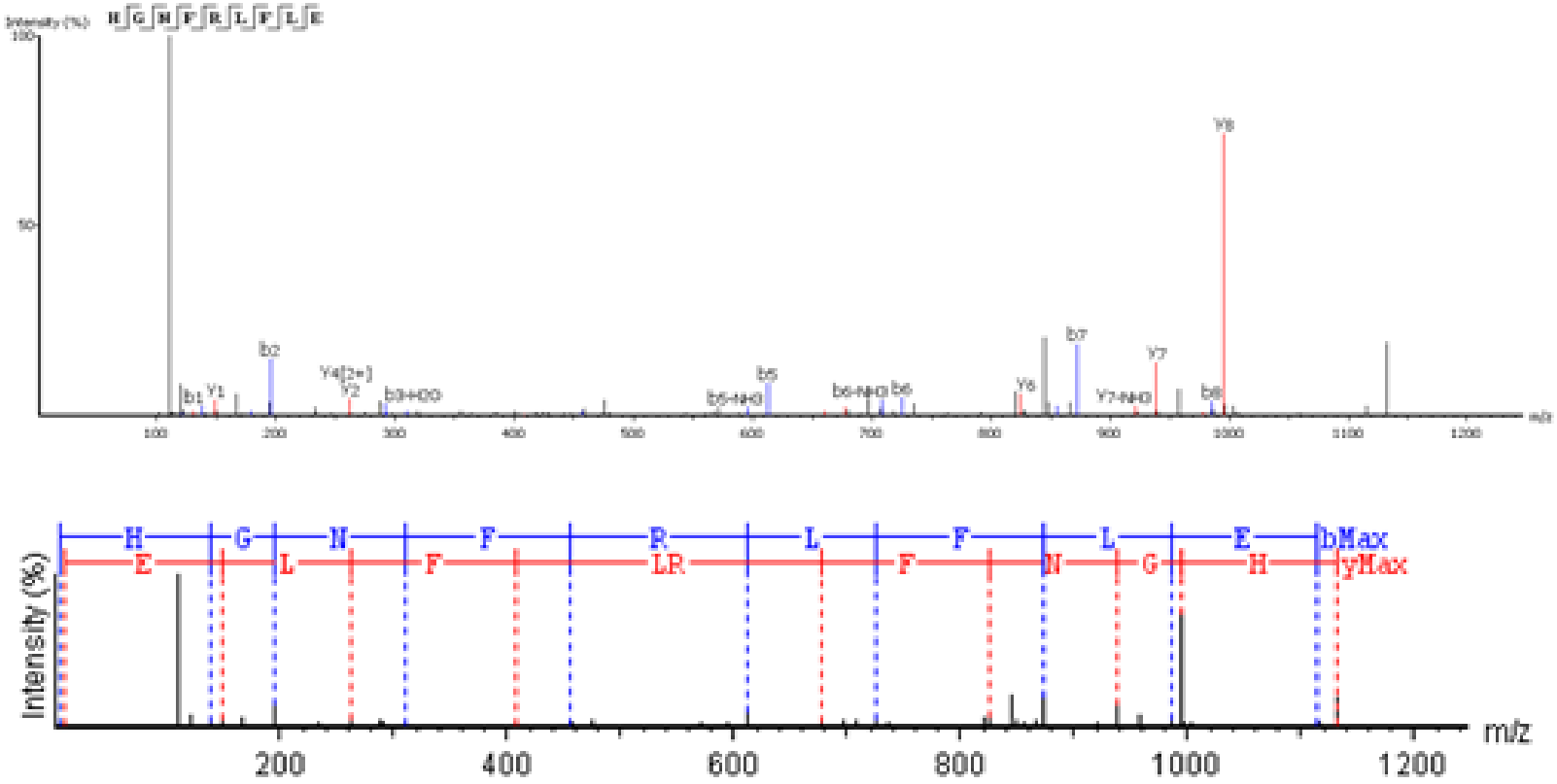
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1980 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



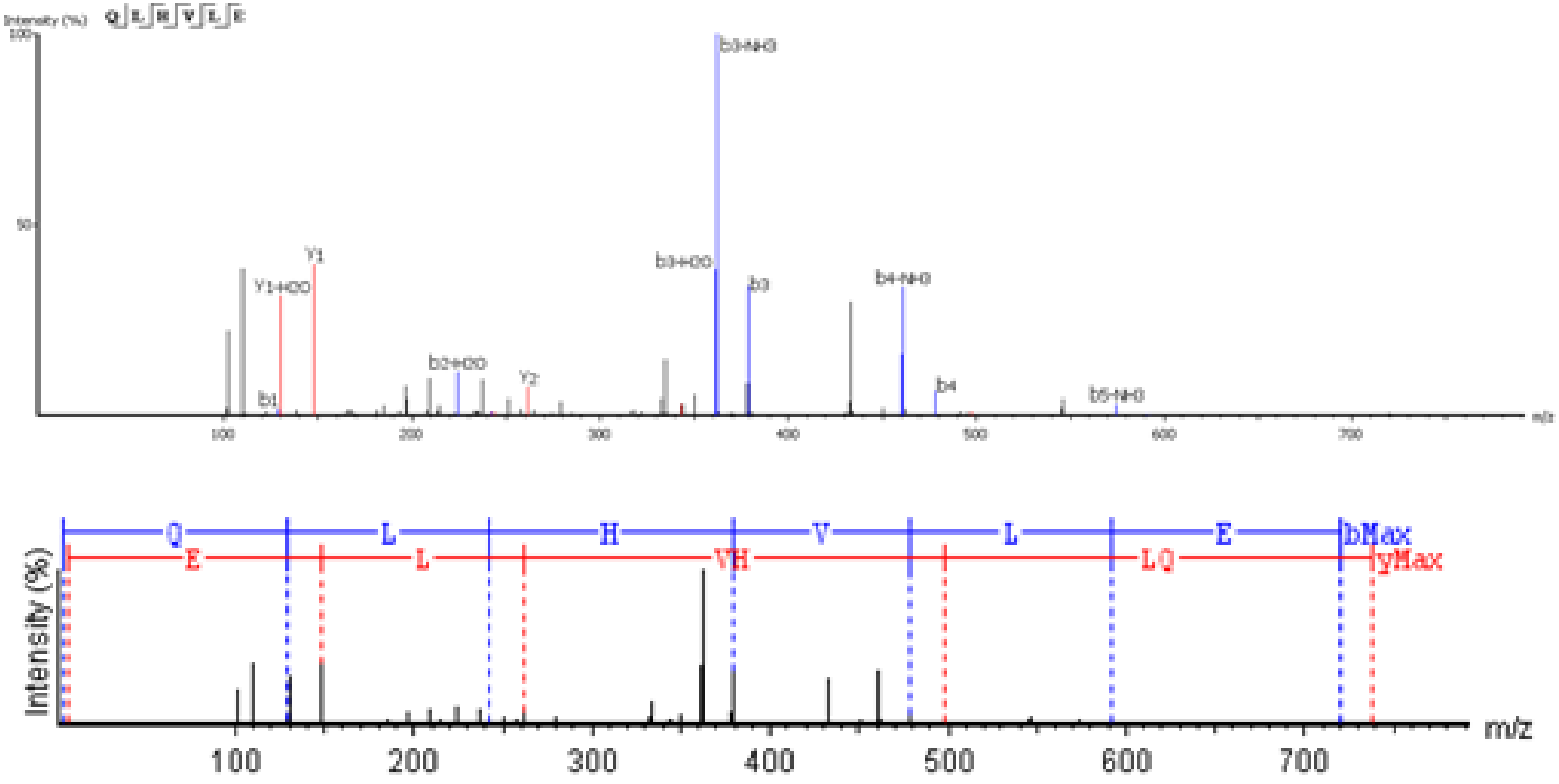
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1967 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



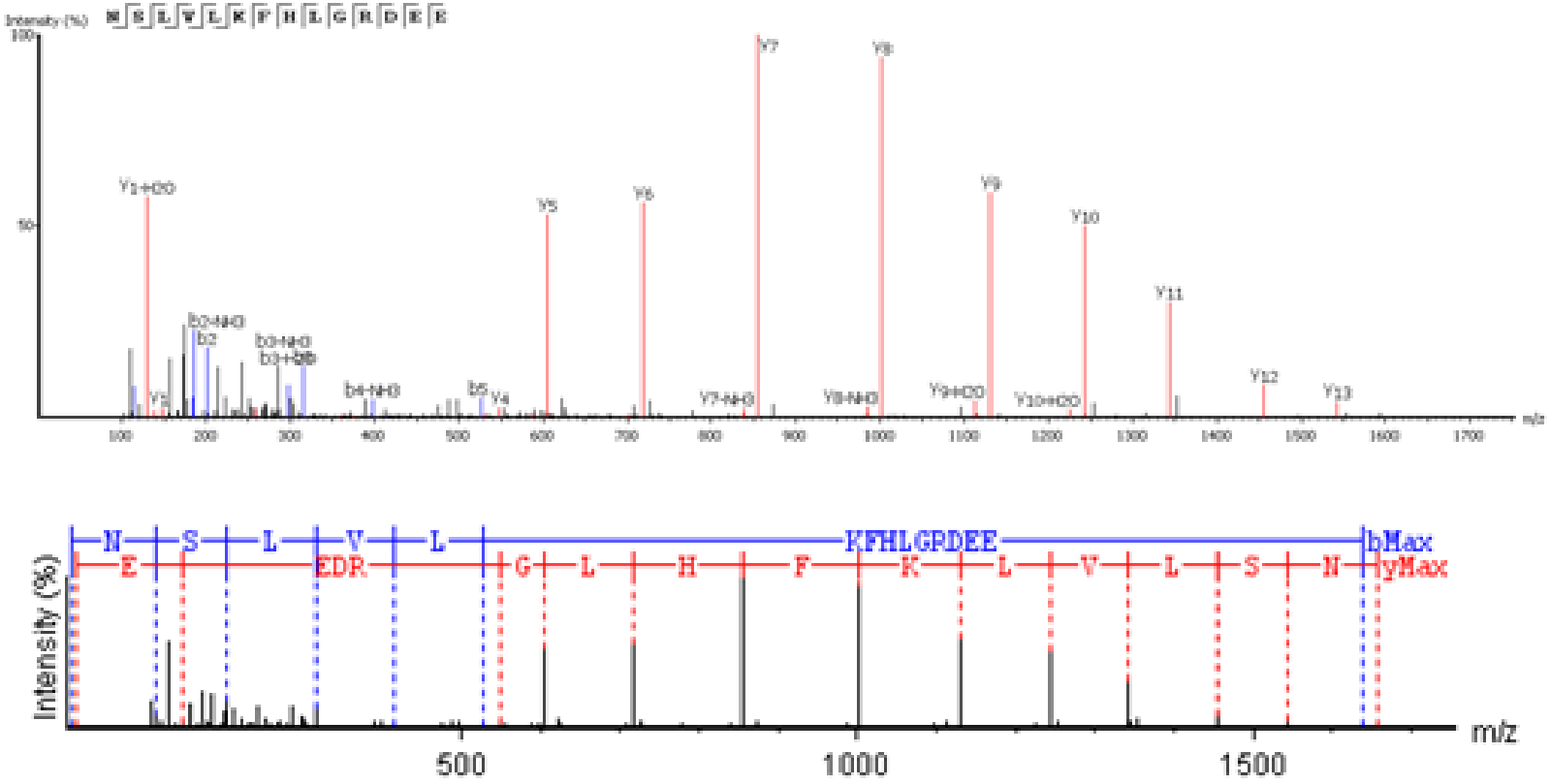
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1132 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



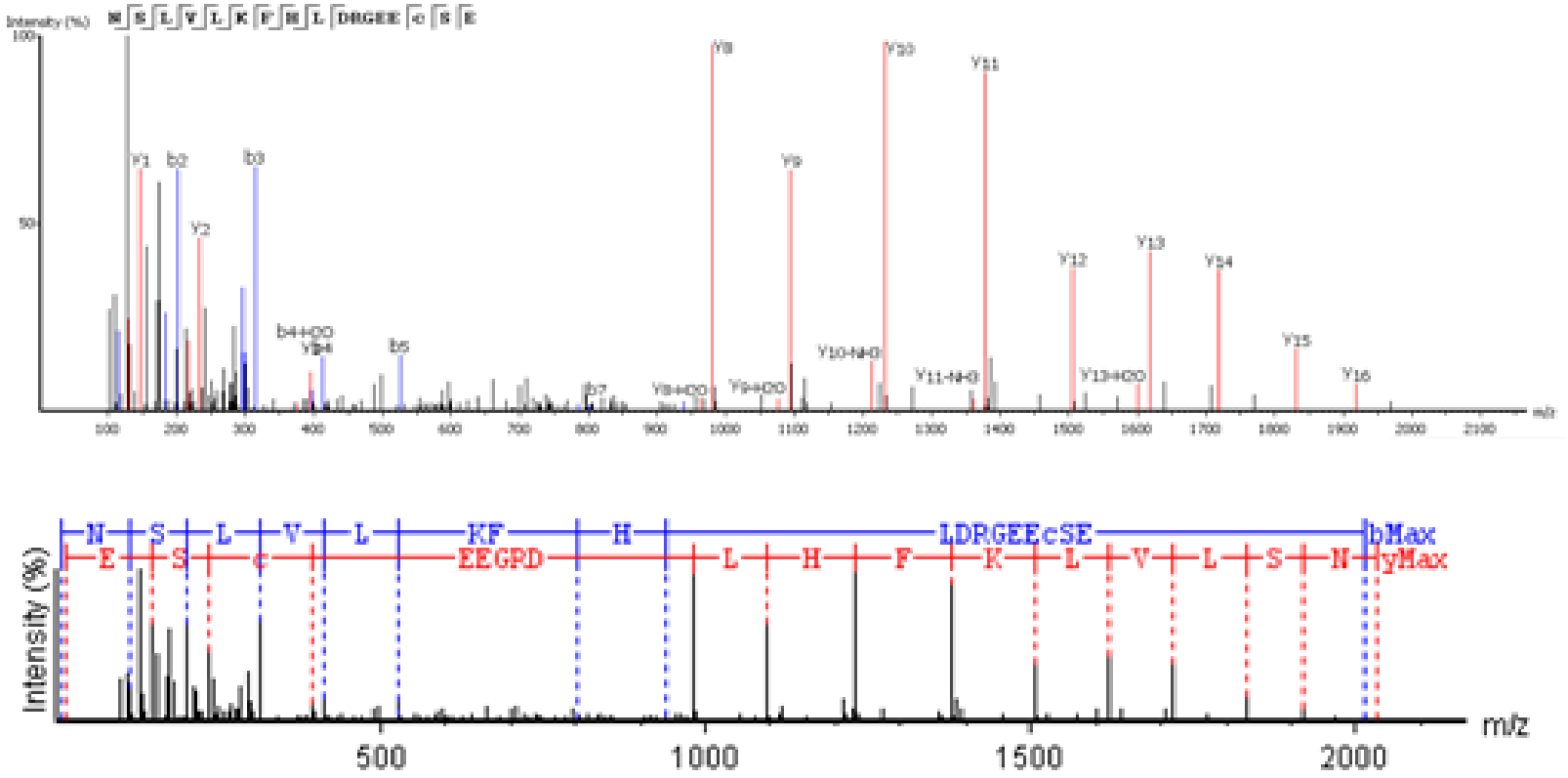
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 737 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



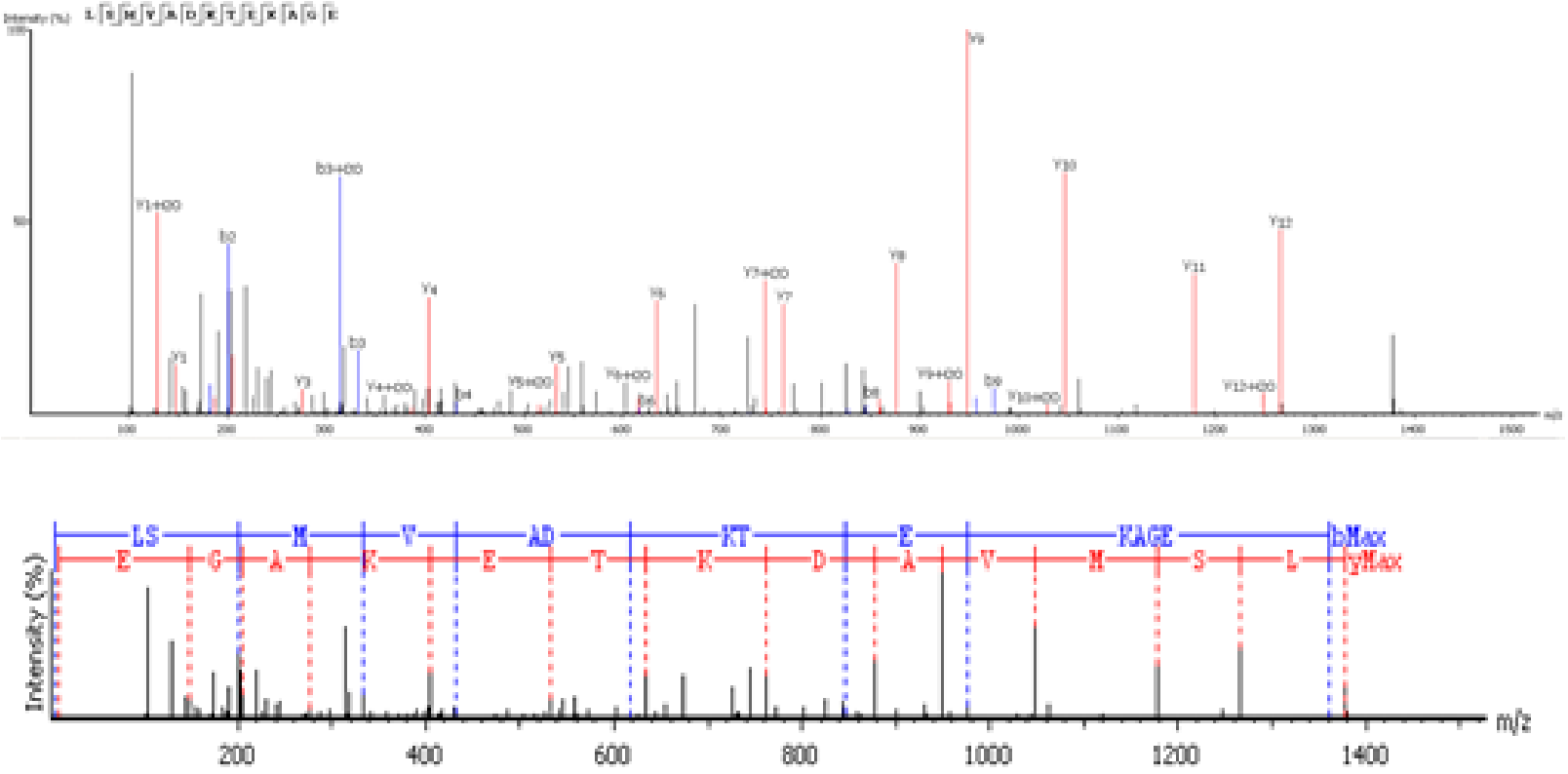
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1656 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



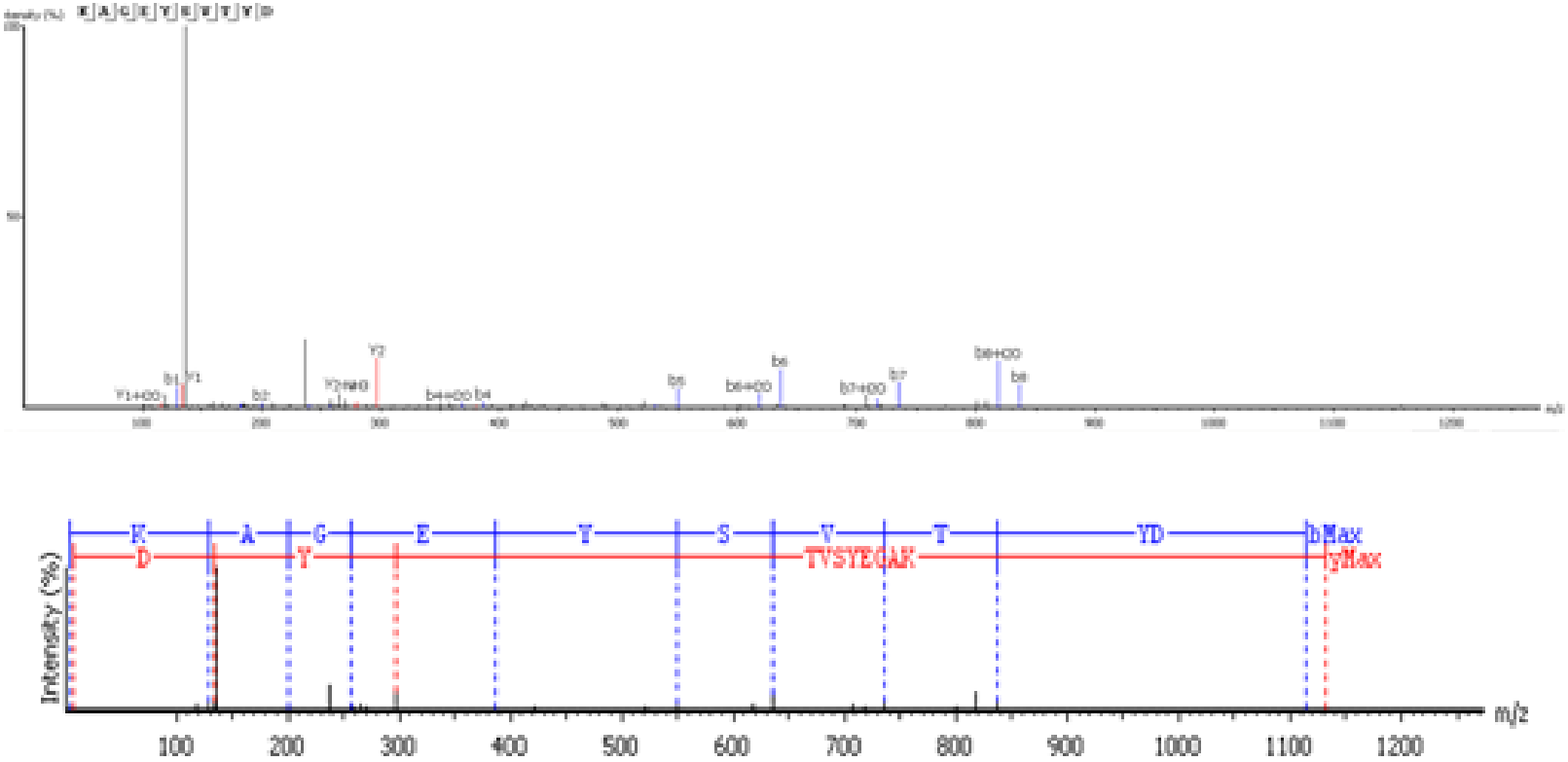
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 2032 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



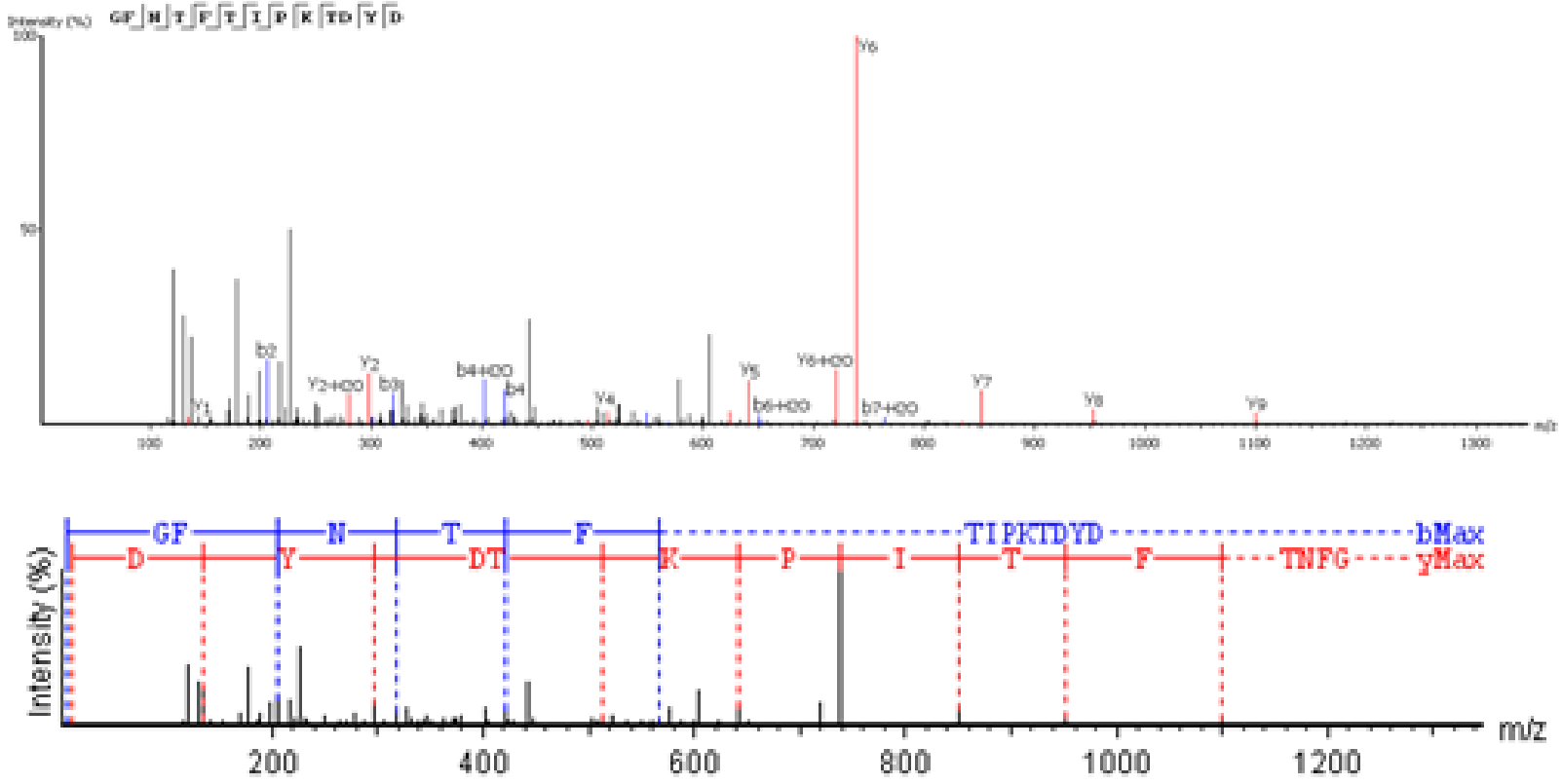
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1378 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



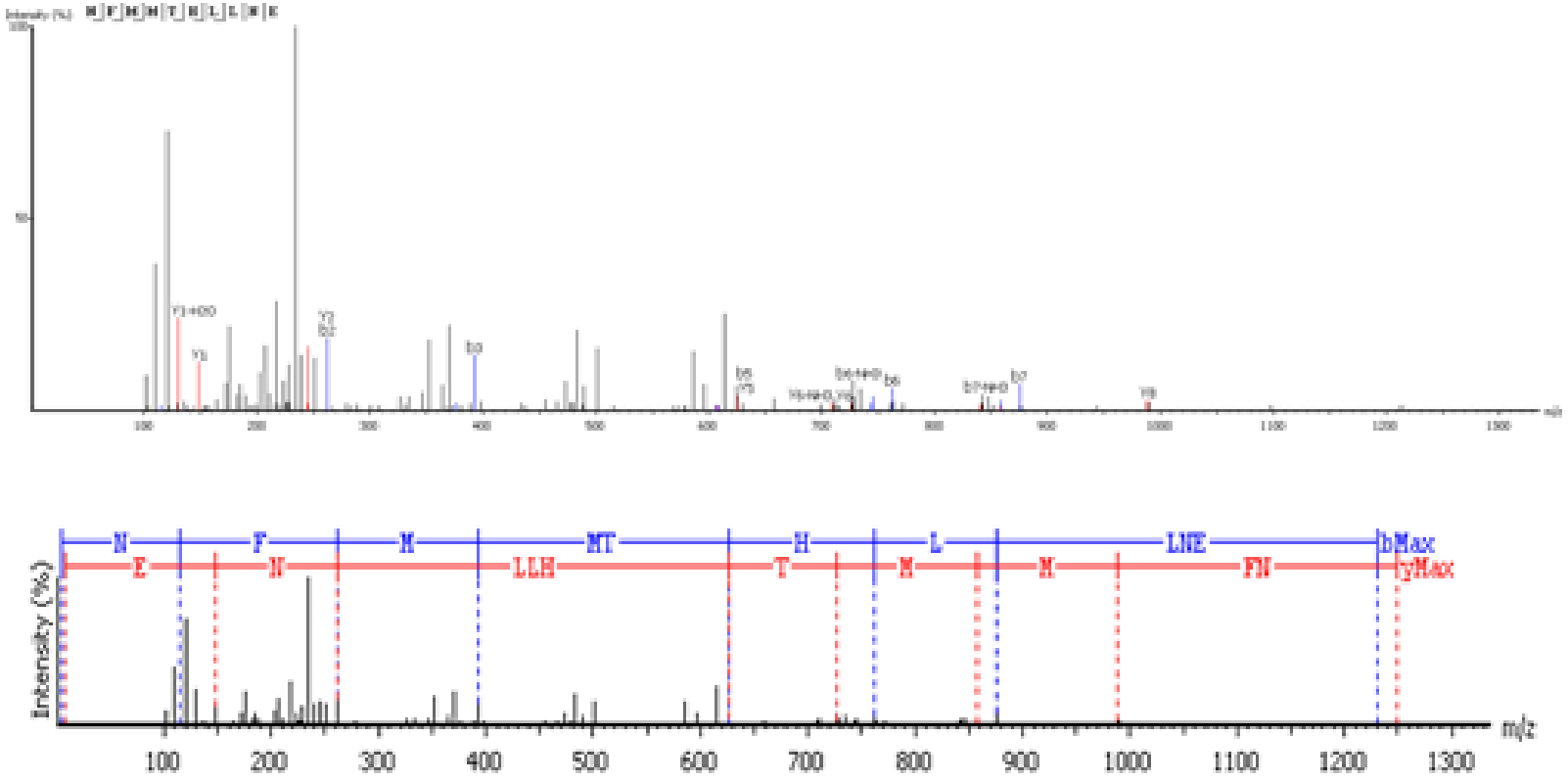
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1132 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



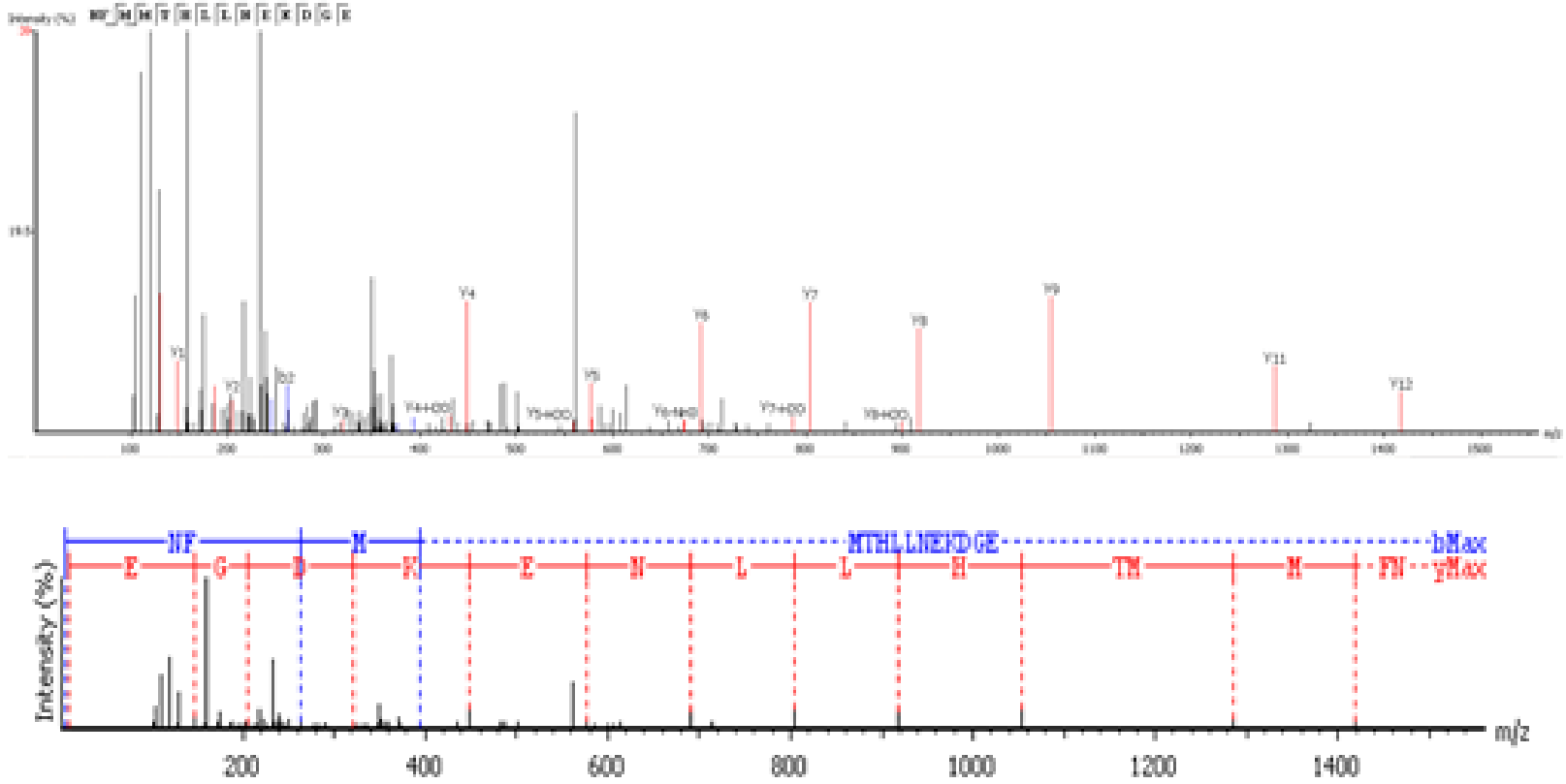
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1518 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



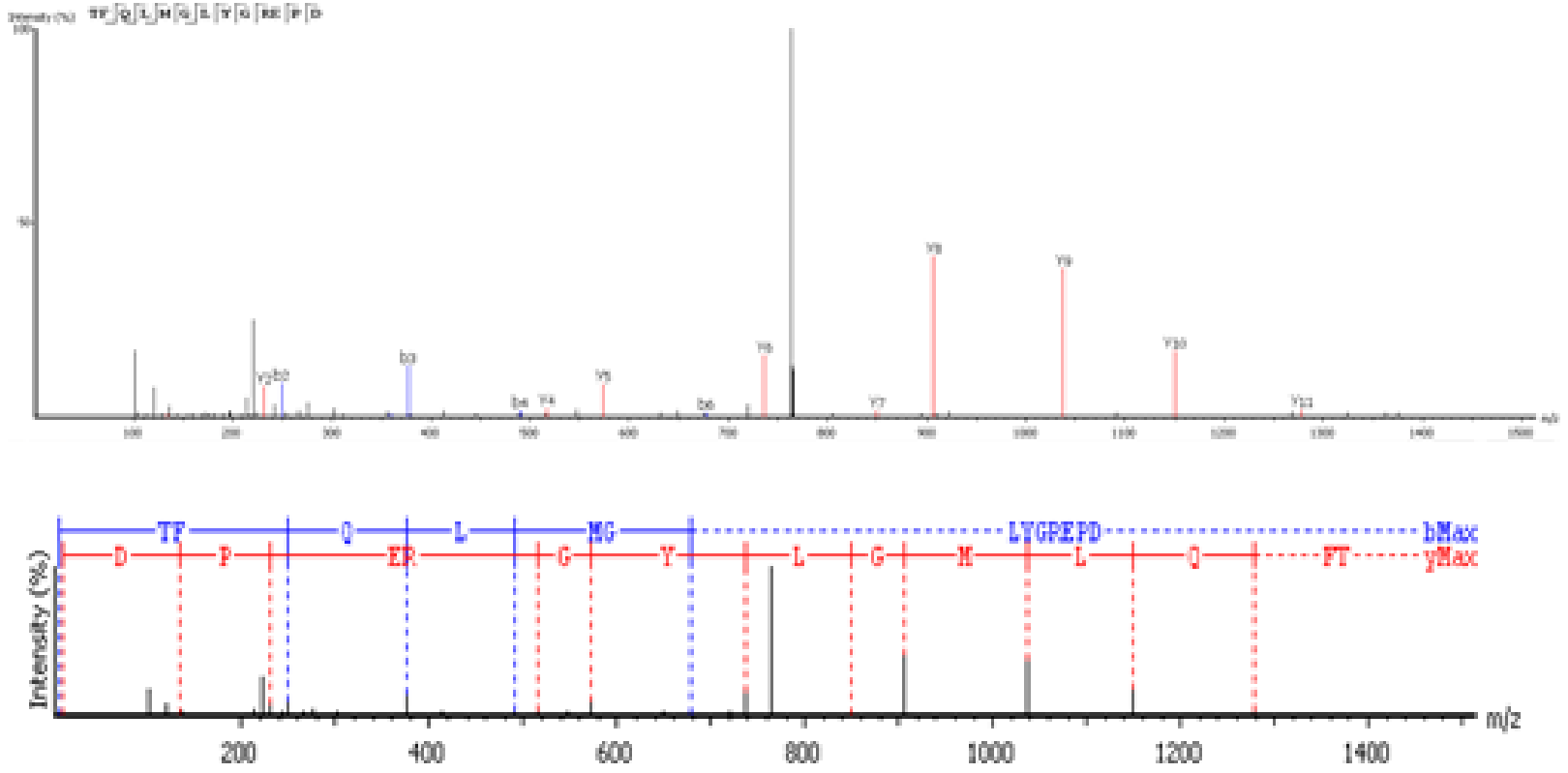
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1249 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



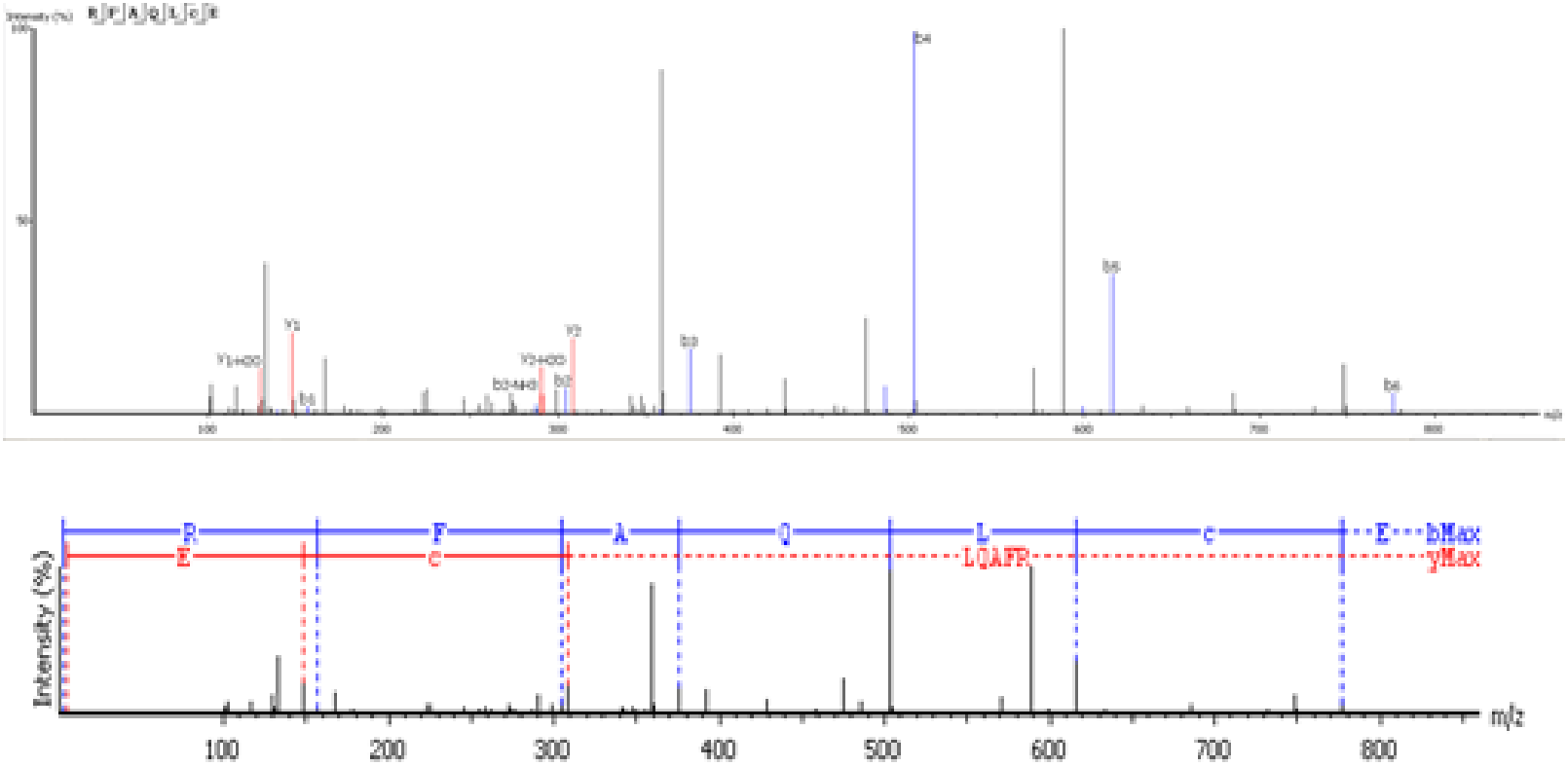
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1678 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



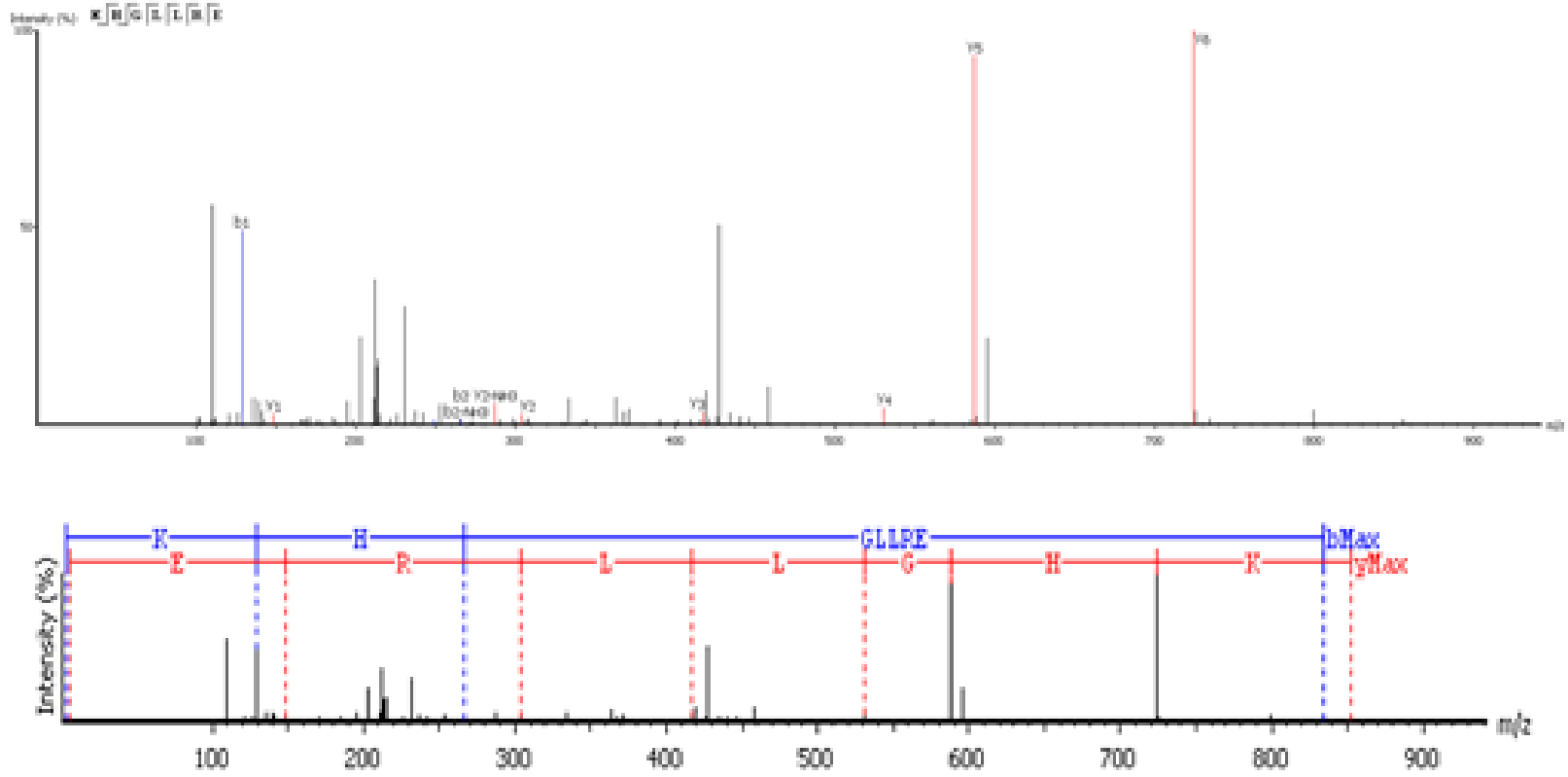
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1526 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



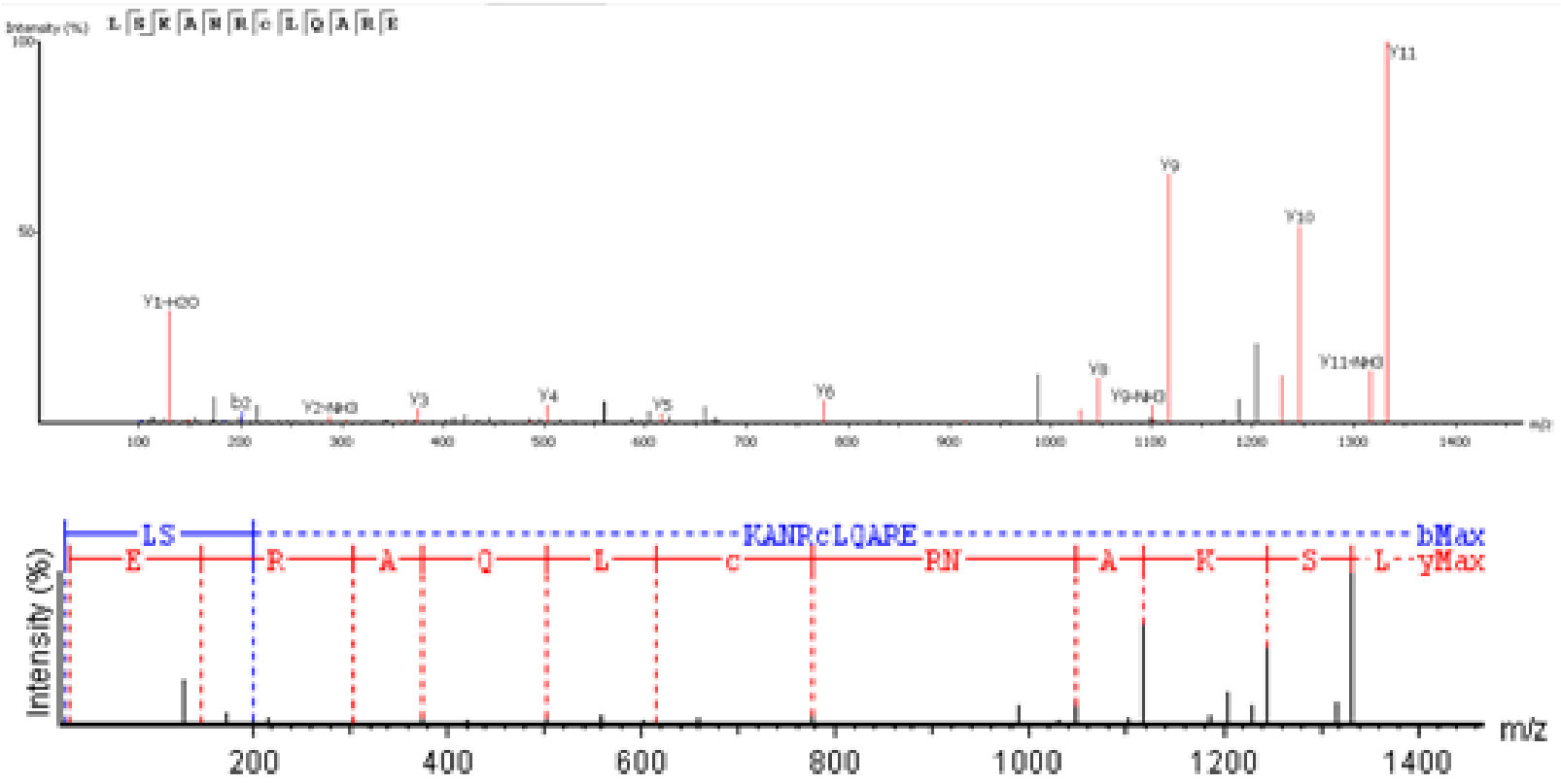
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 922 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



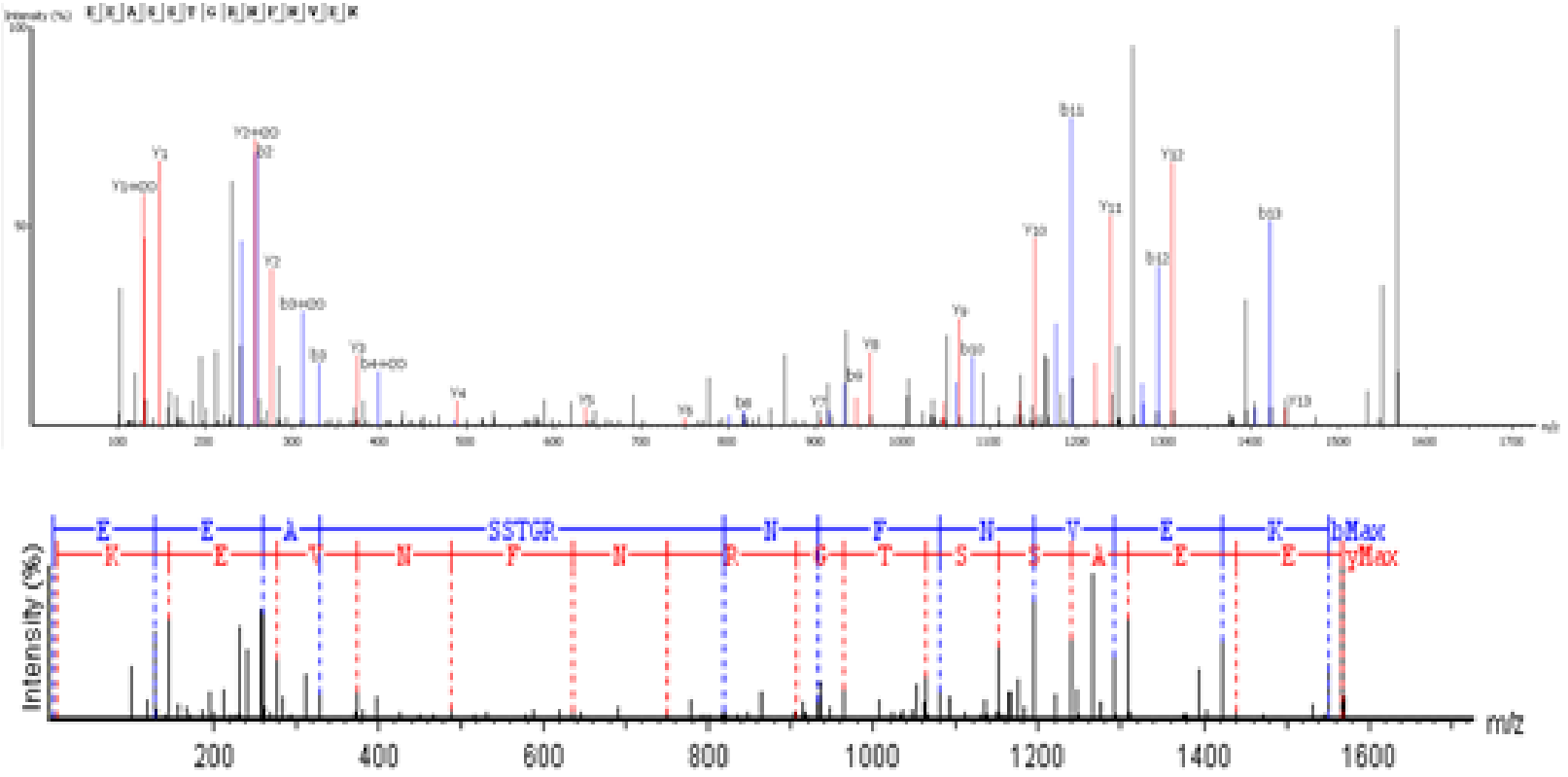
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 852 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



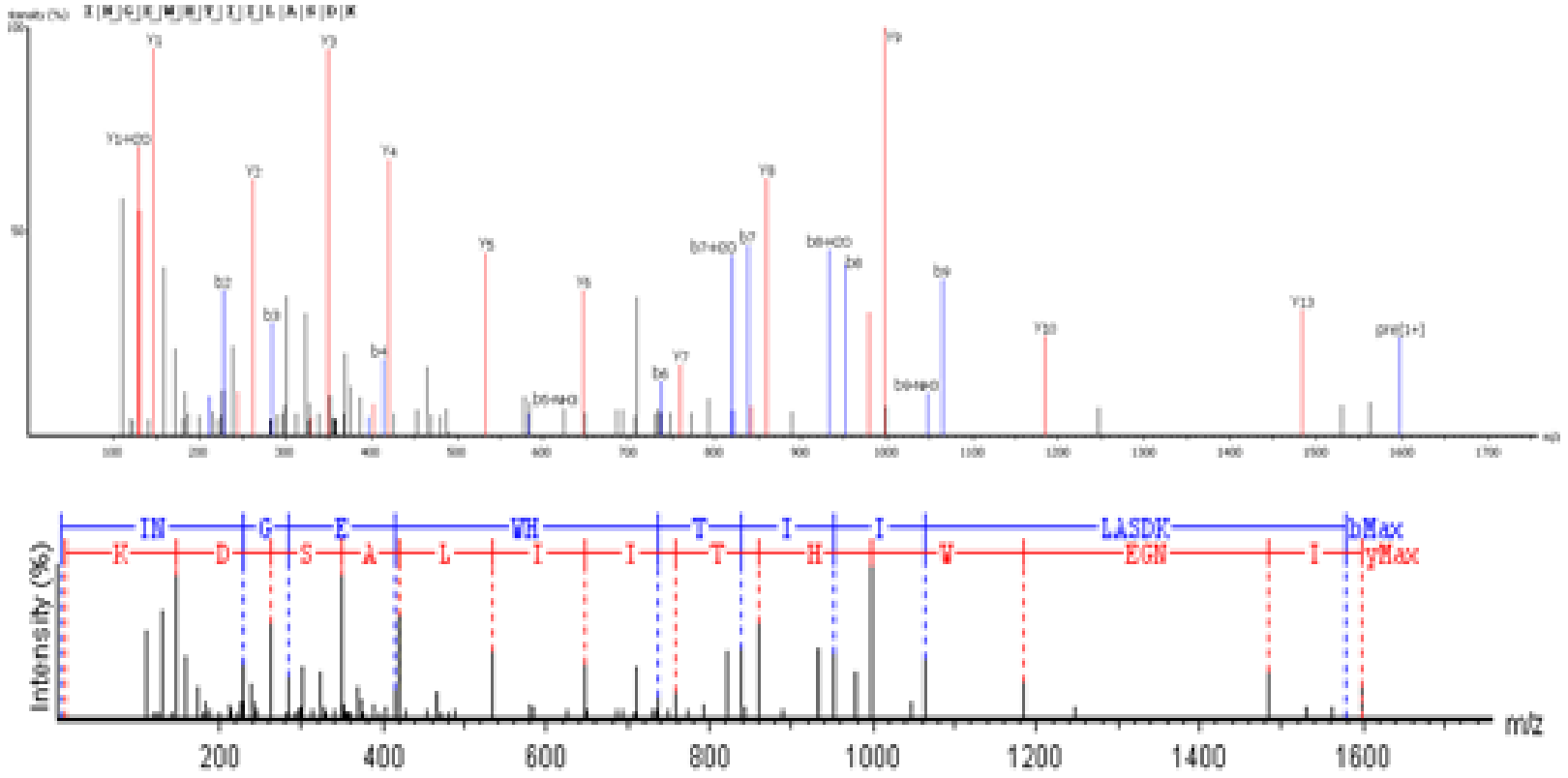
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1445 Da of 18762 Da MUP.

Fraction 3 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



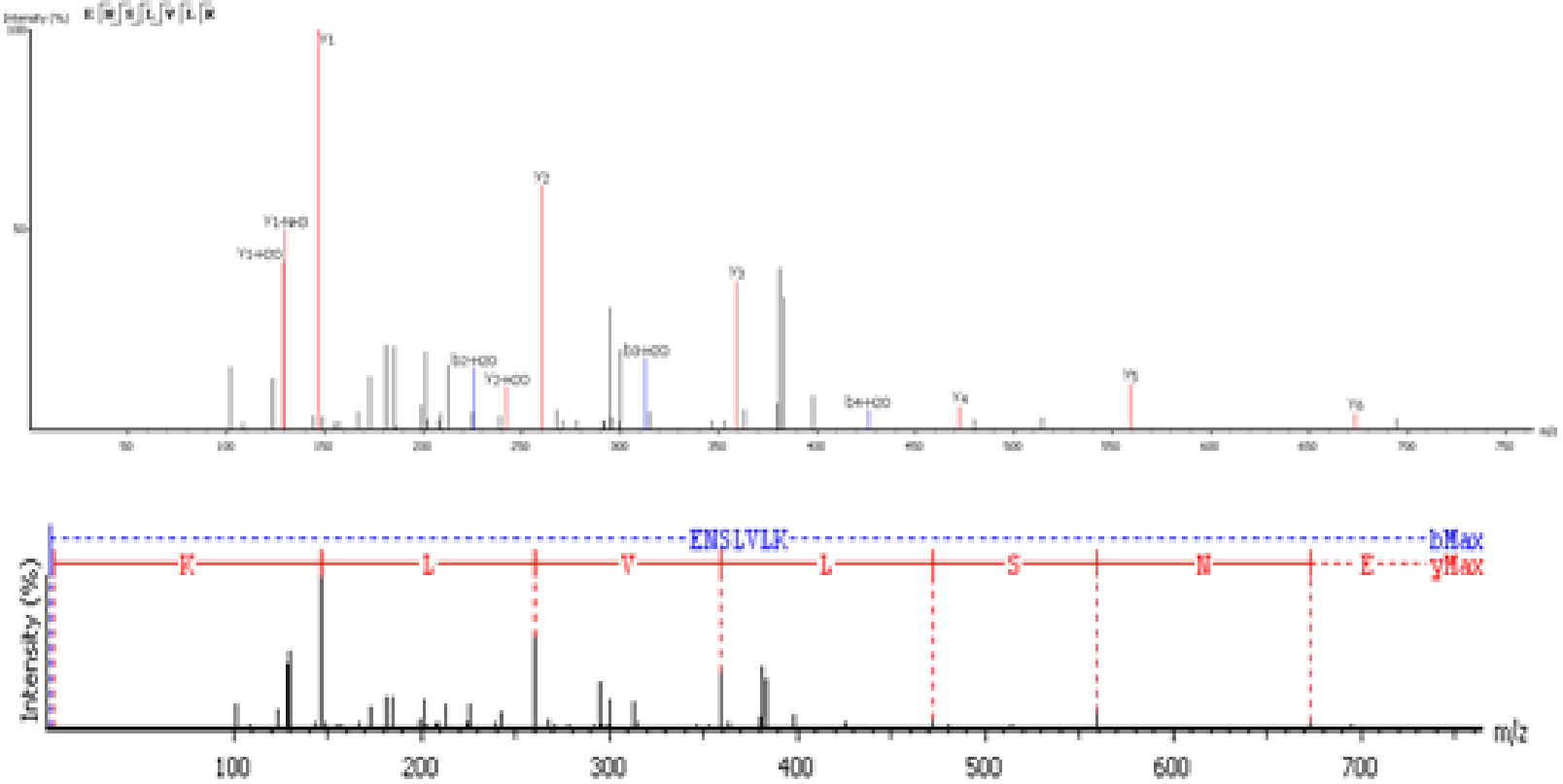
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1567 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



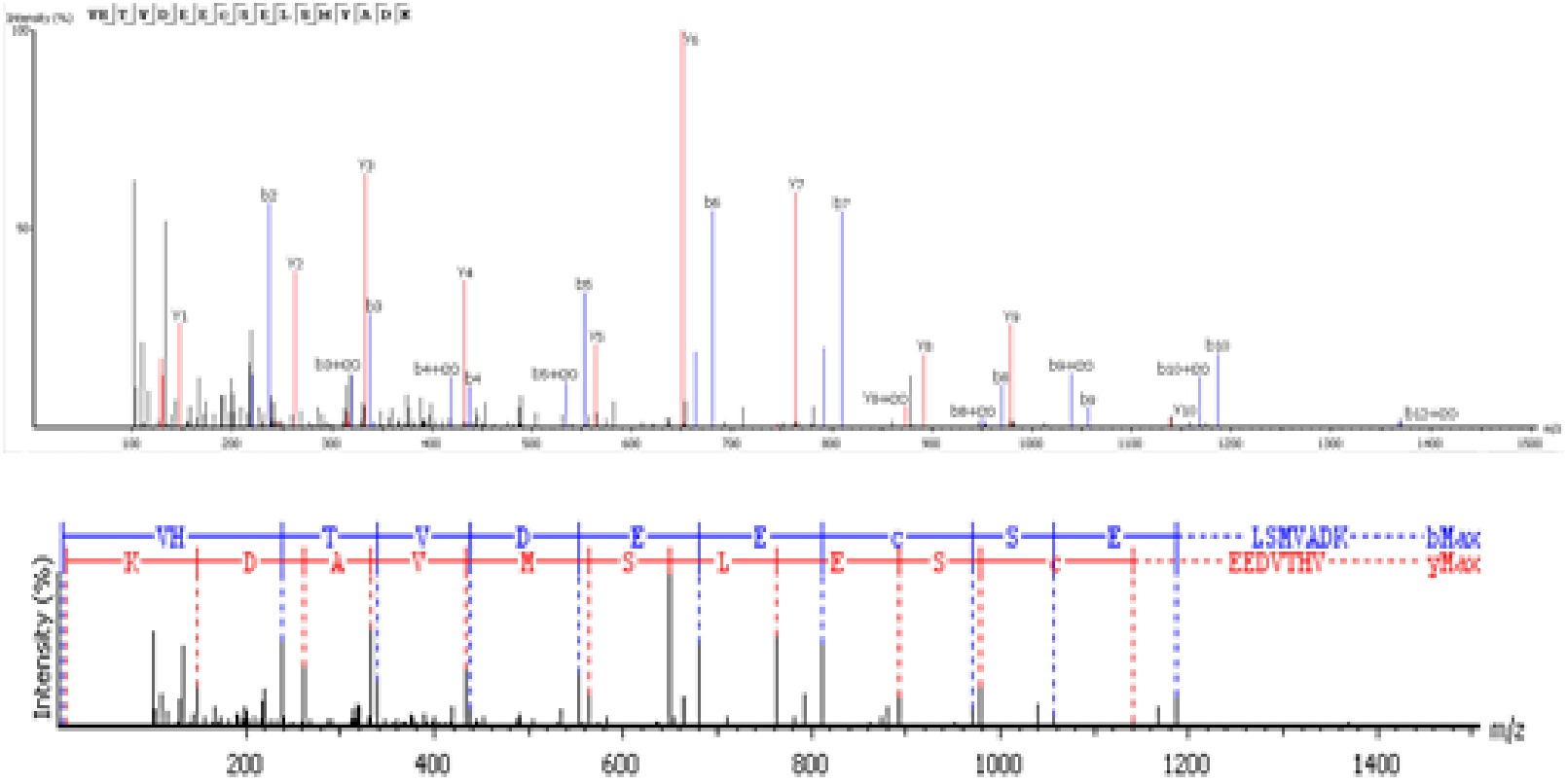
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1597 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



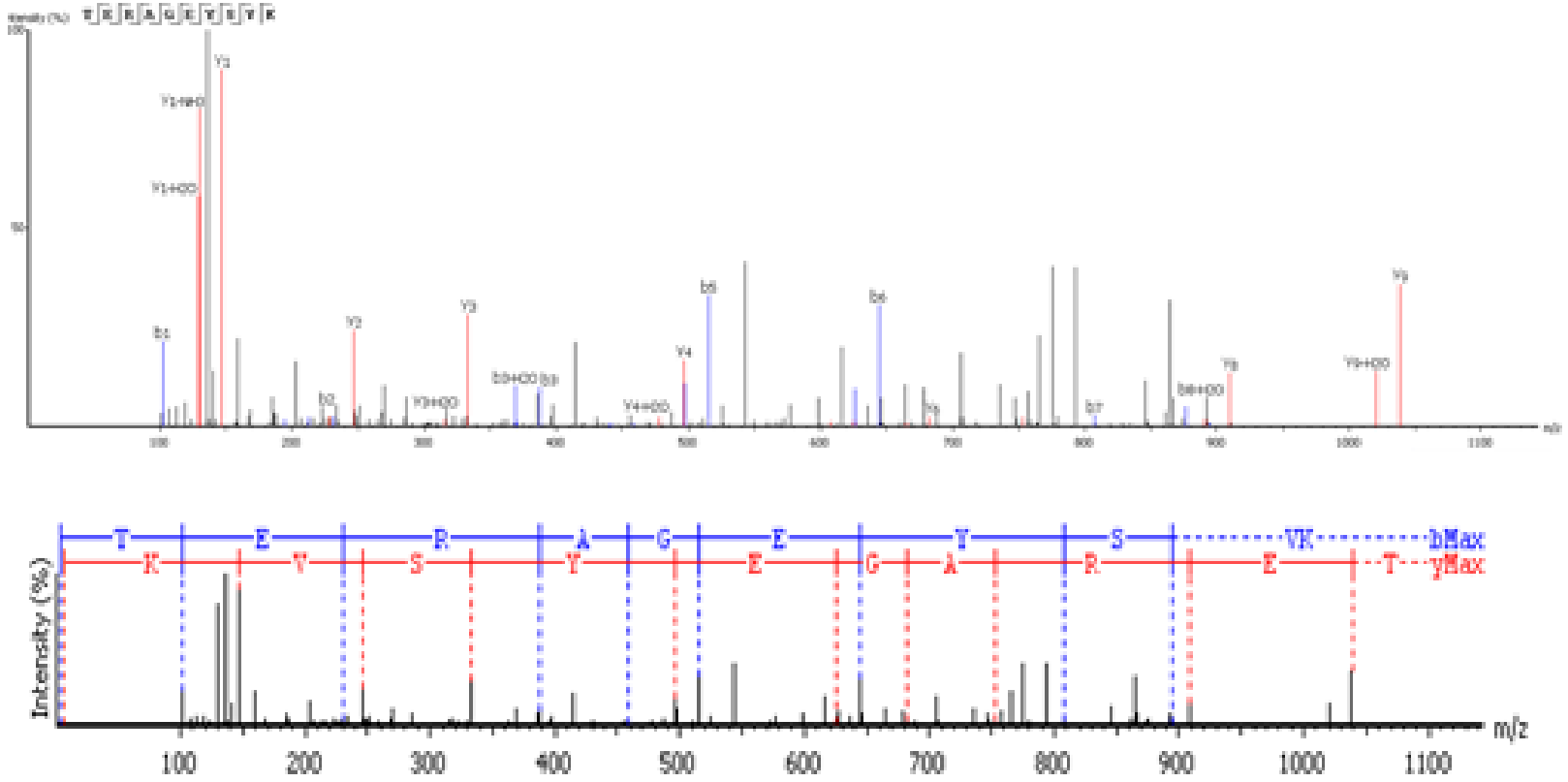
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 801 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



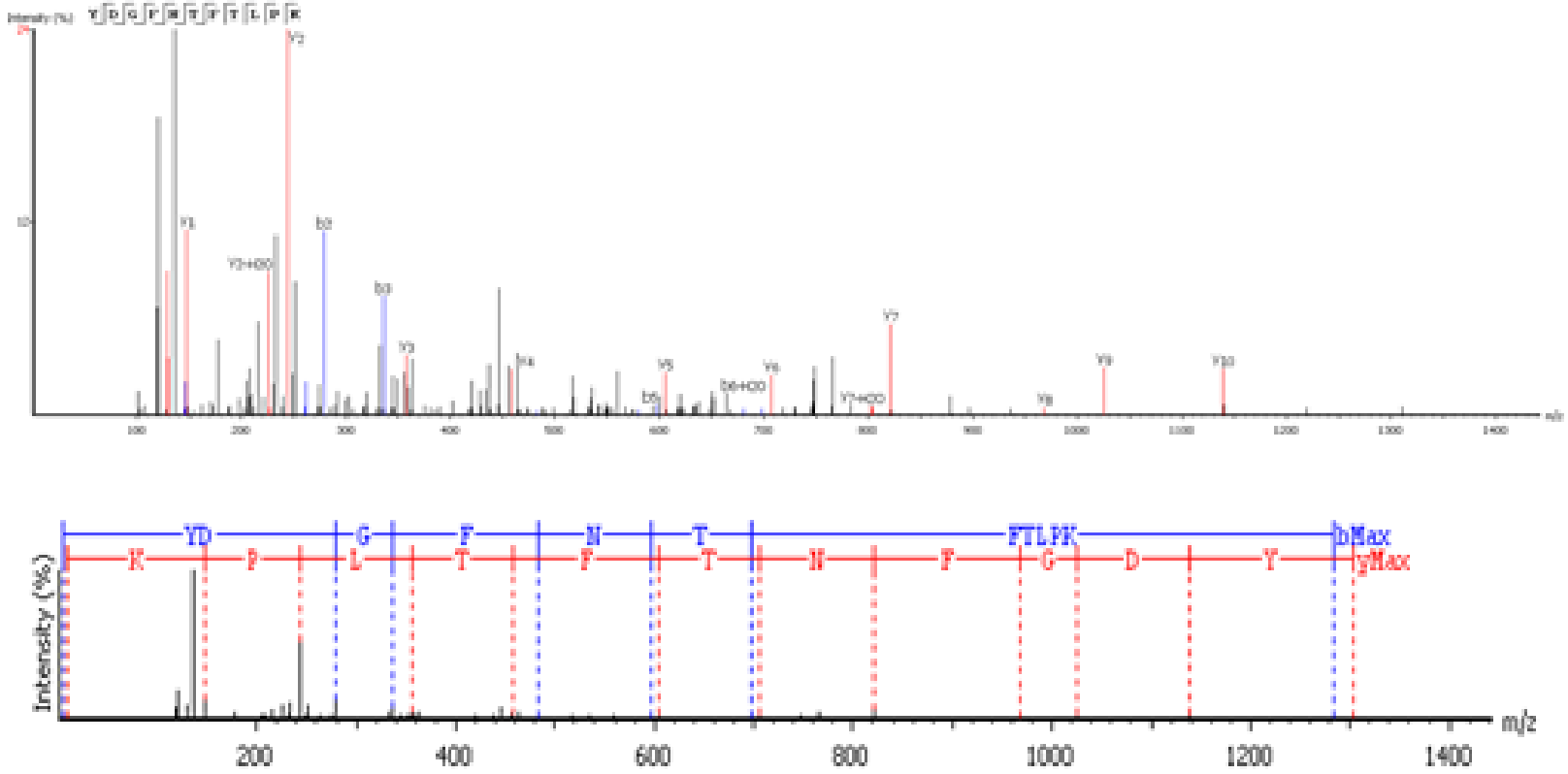
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1948 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



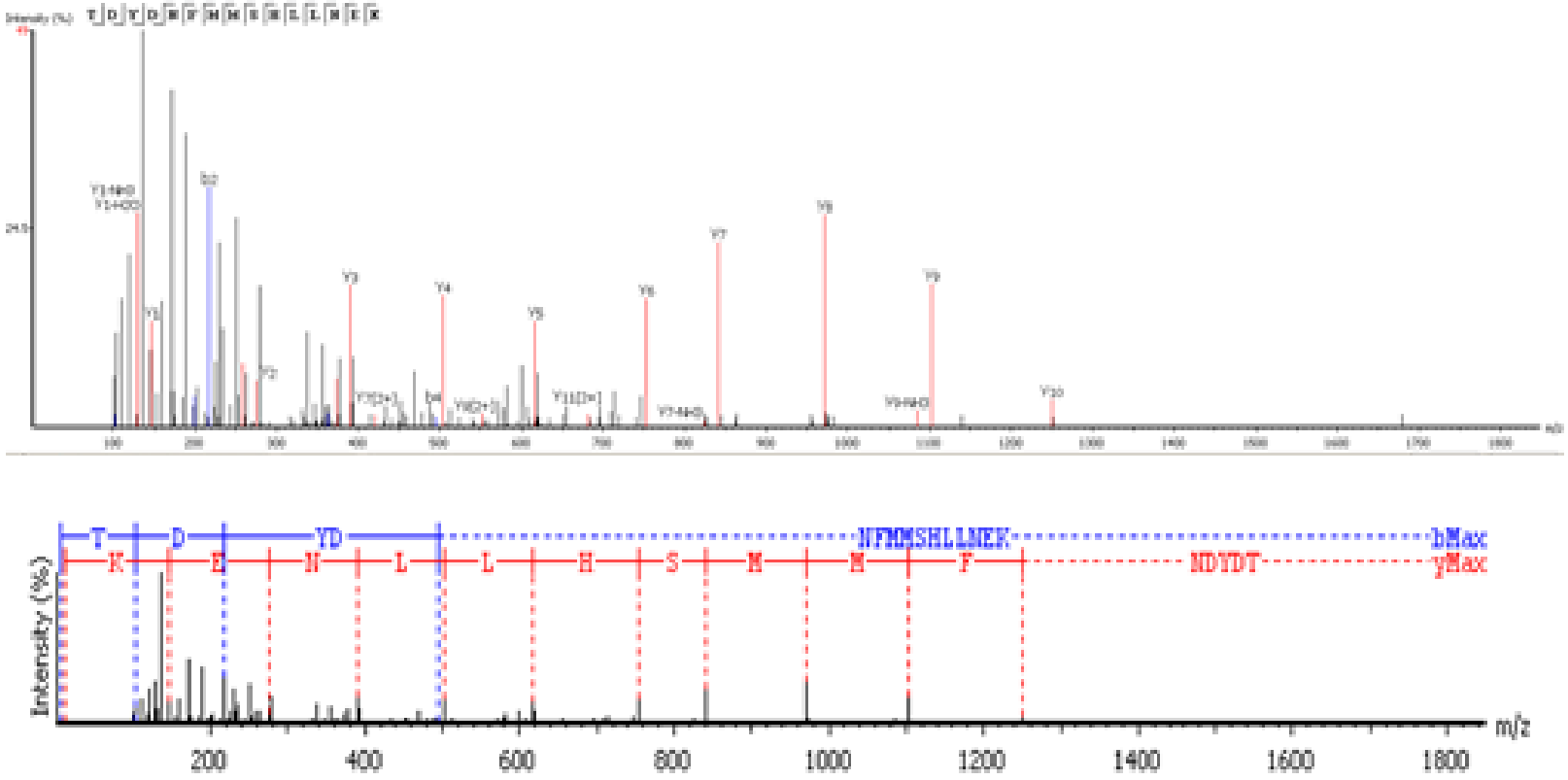
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1139 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



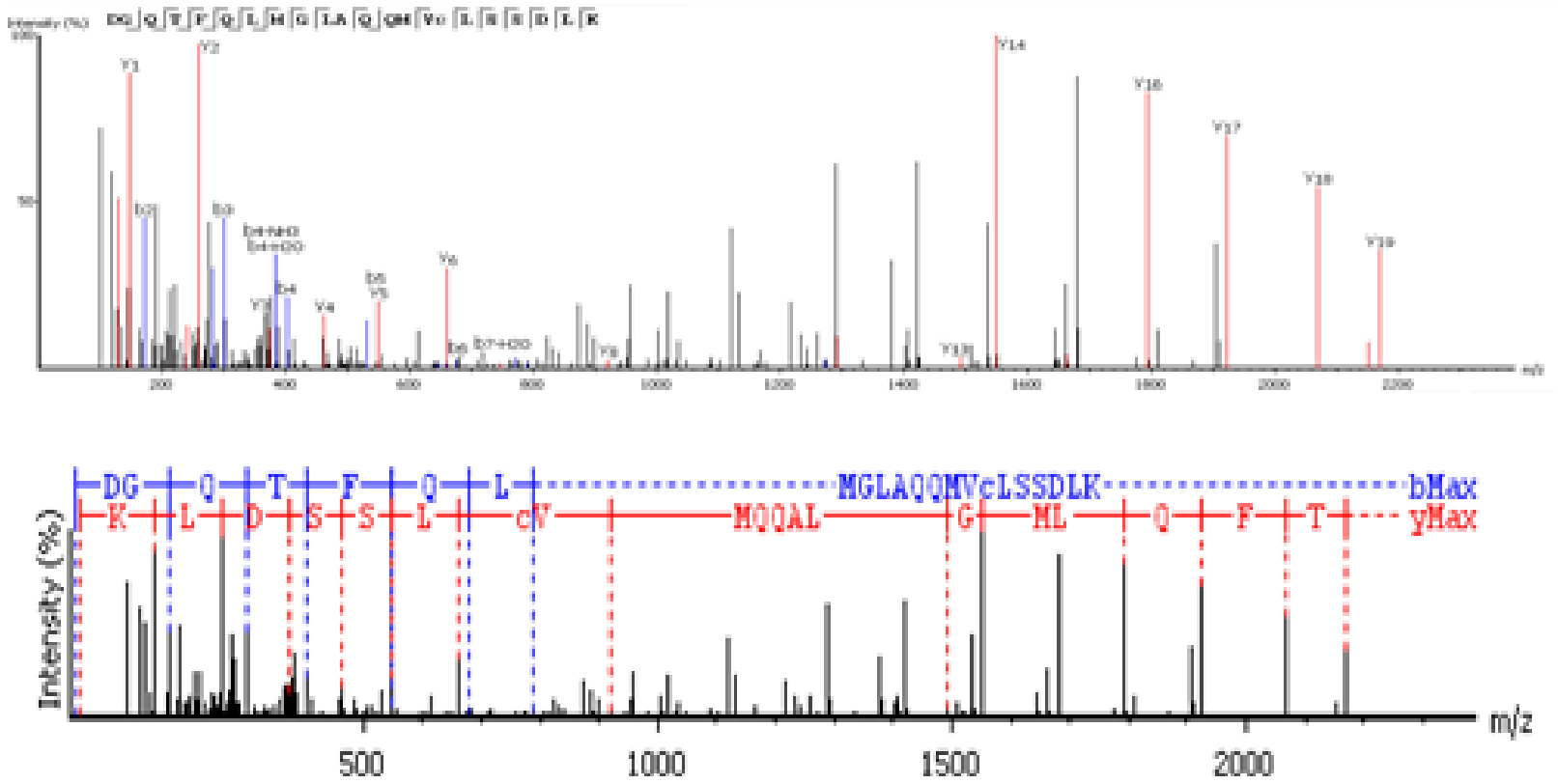
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1302 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



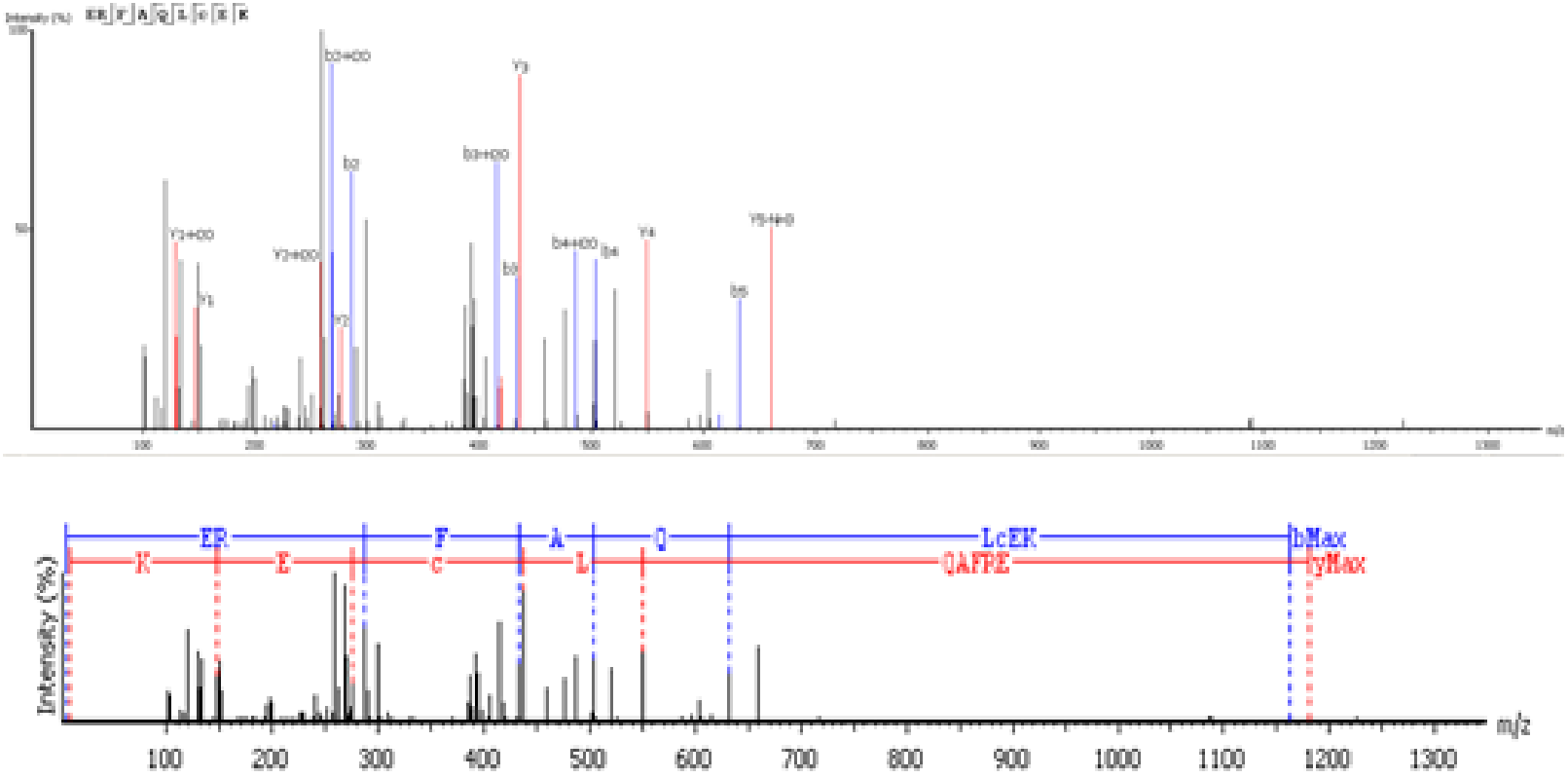
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1857 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



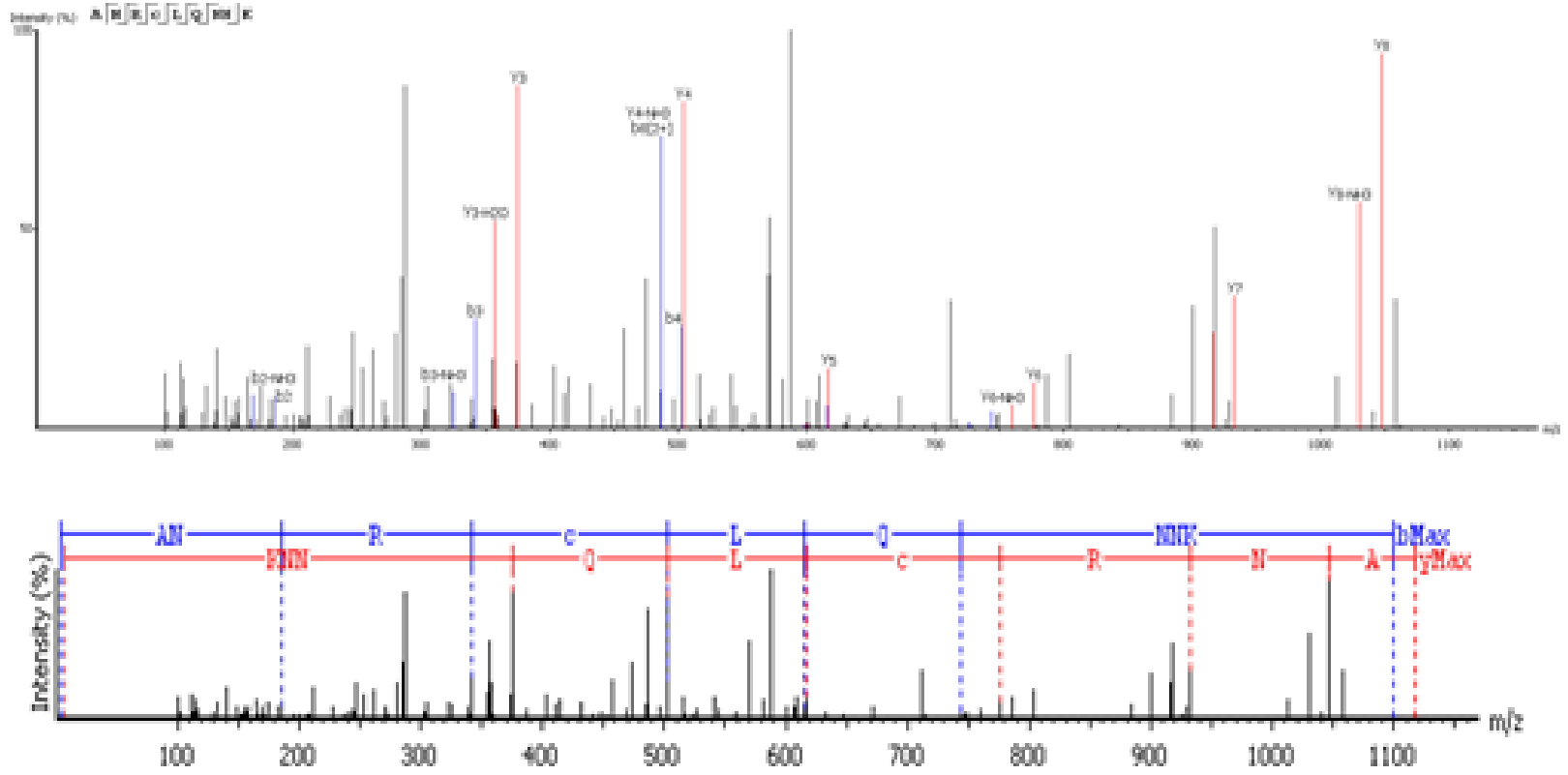
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2469 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



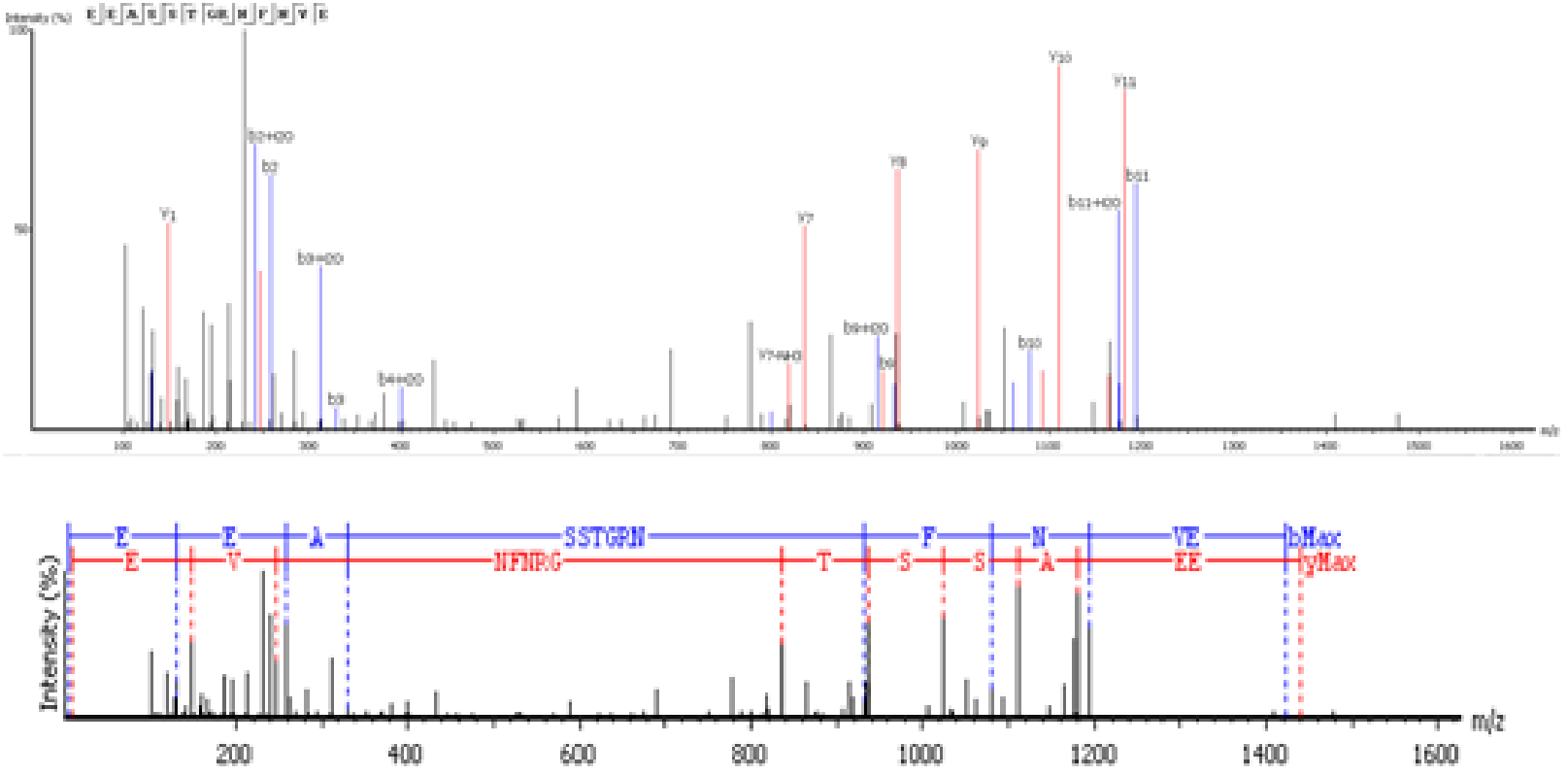
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1180 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



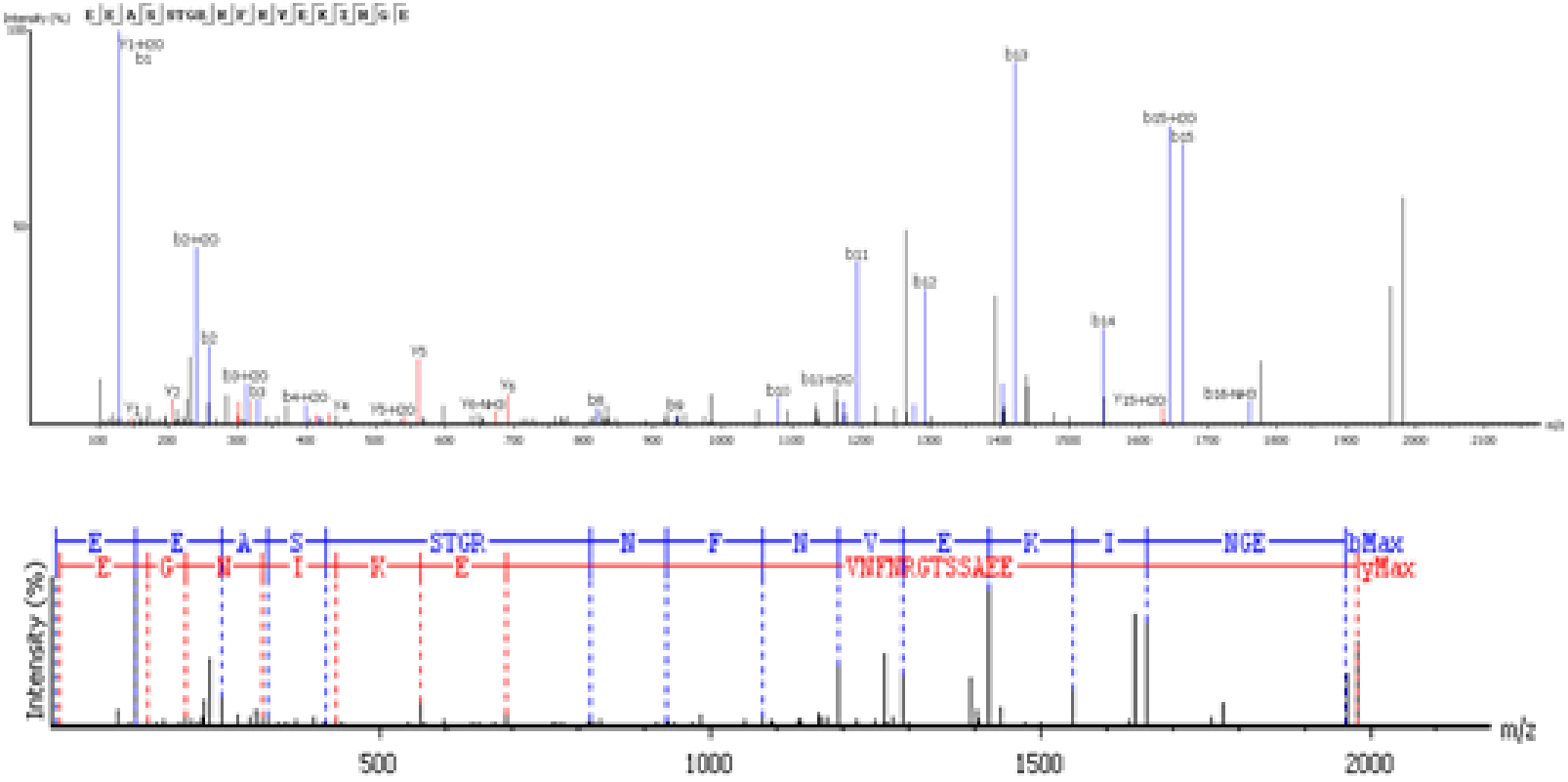
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1117 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



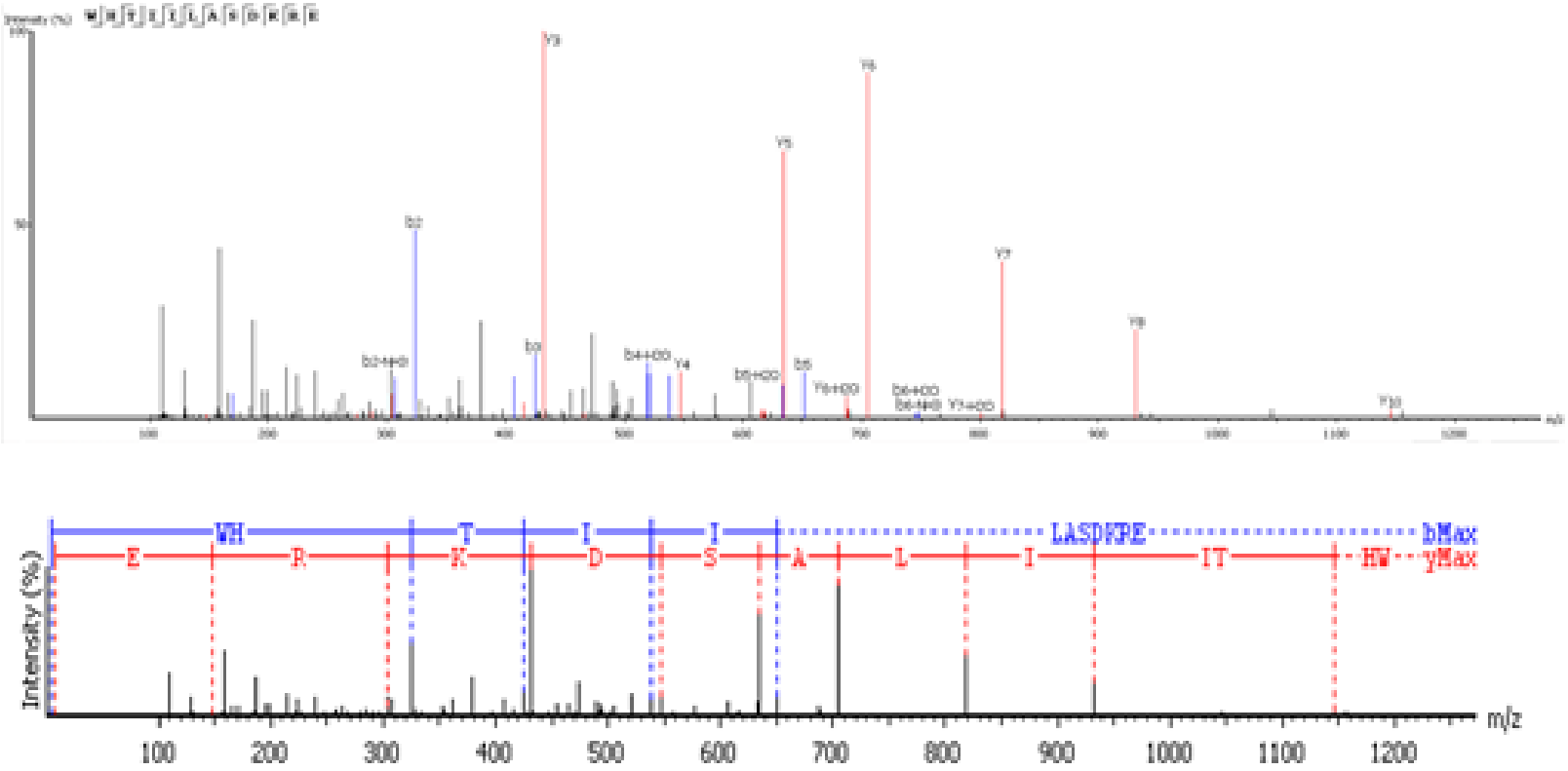
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1439 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



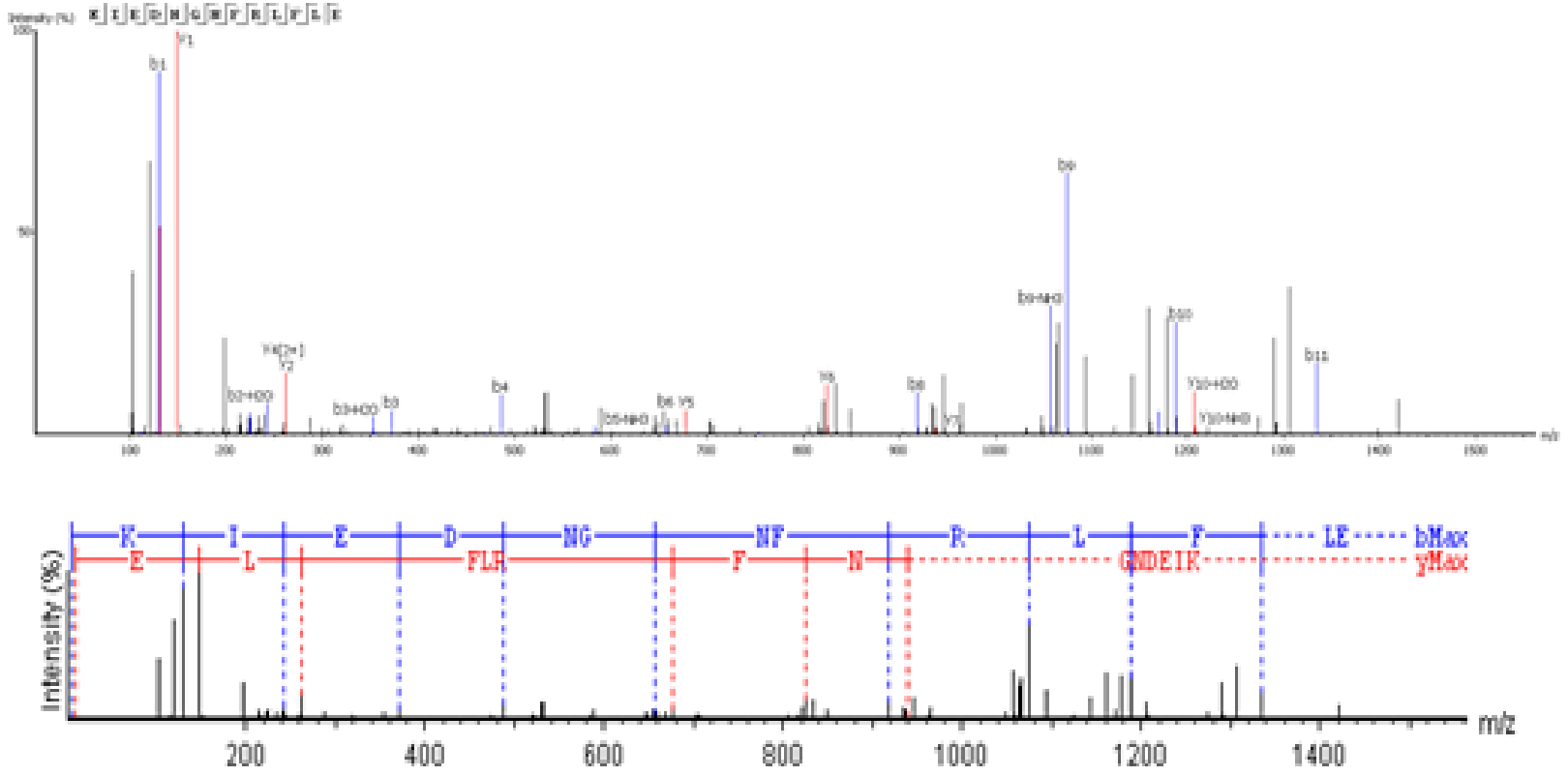
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1980 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



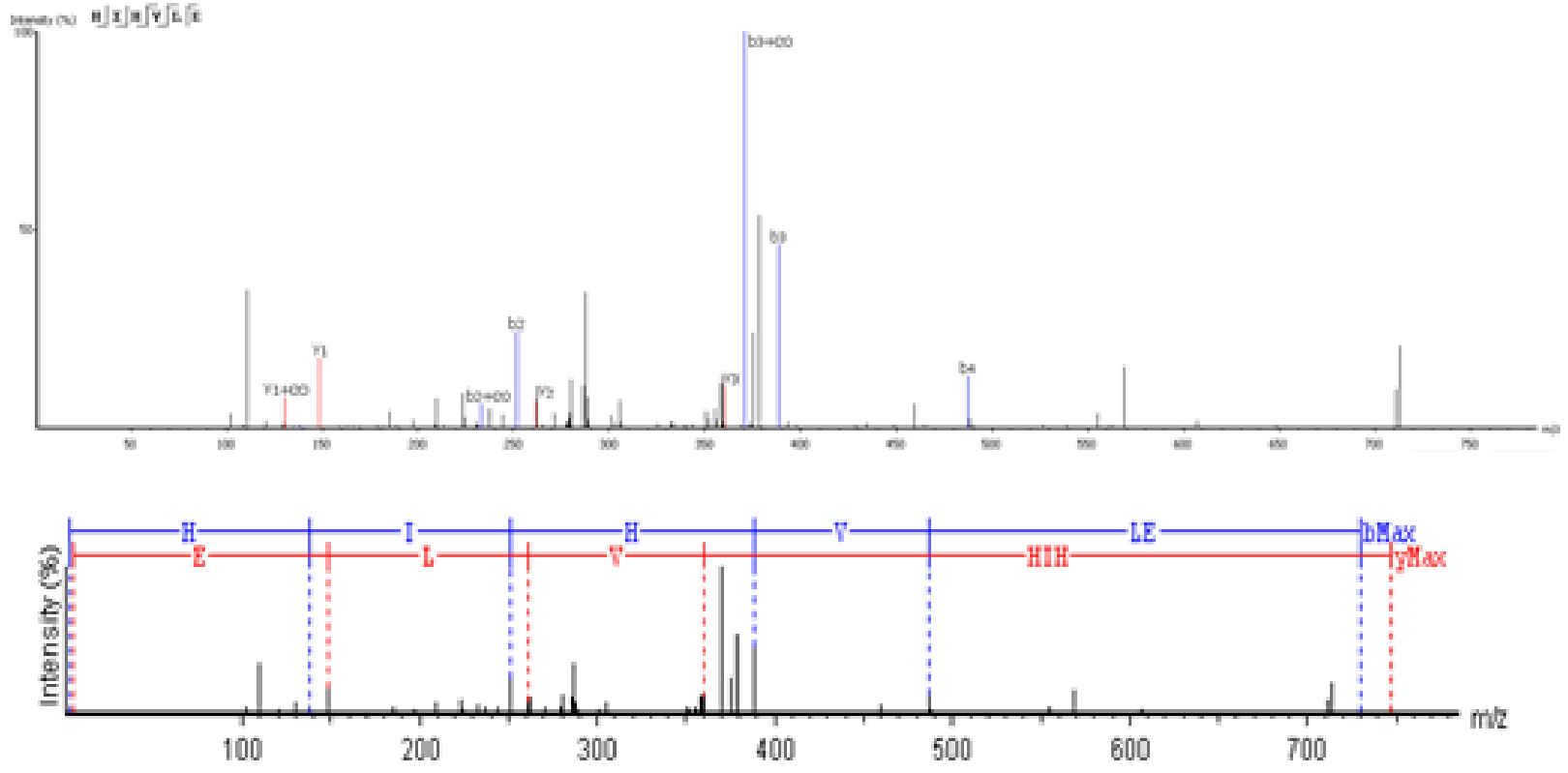
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1468 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



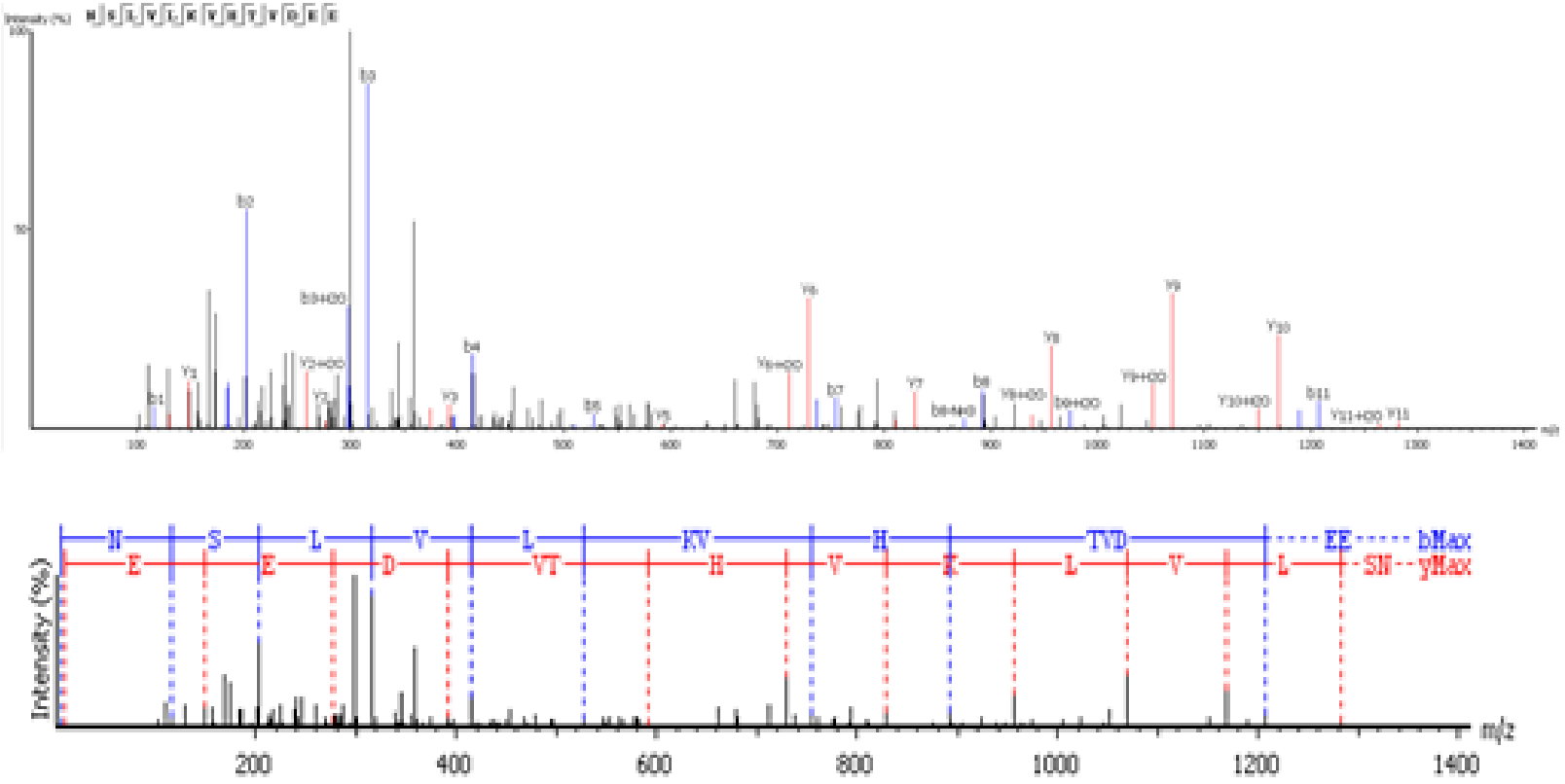
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1594 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



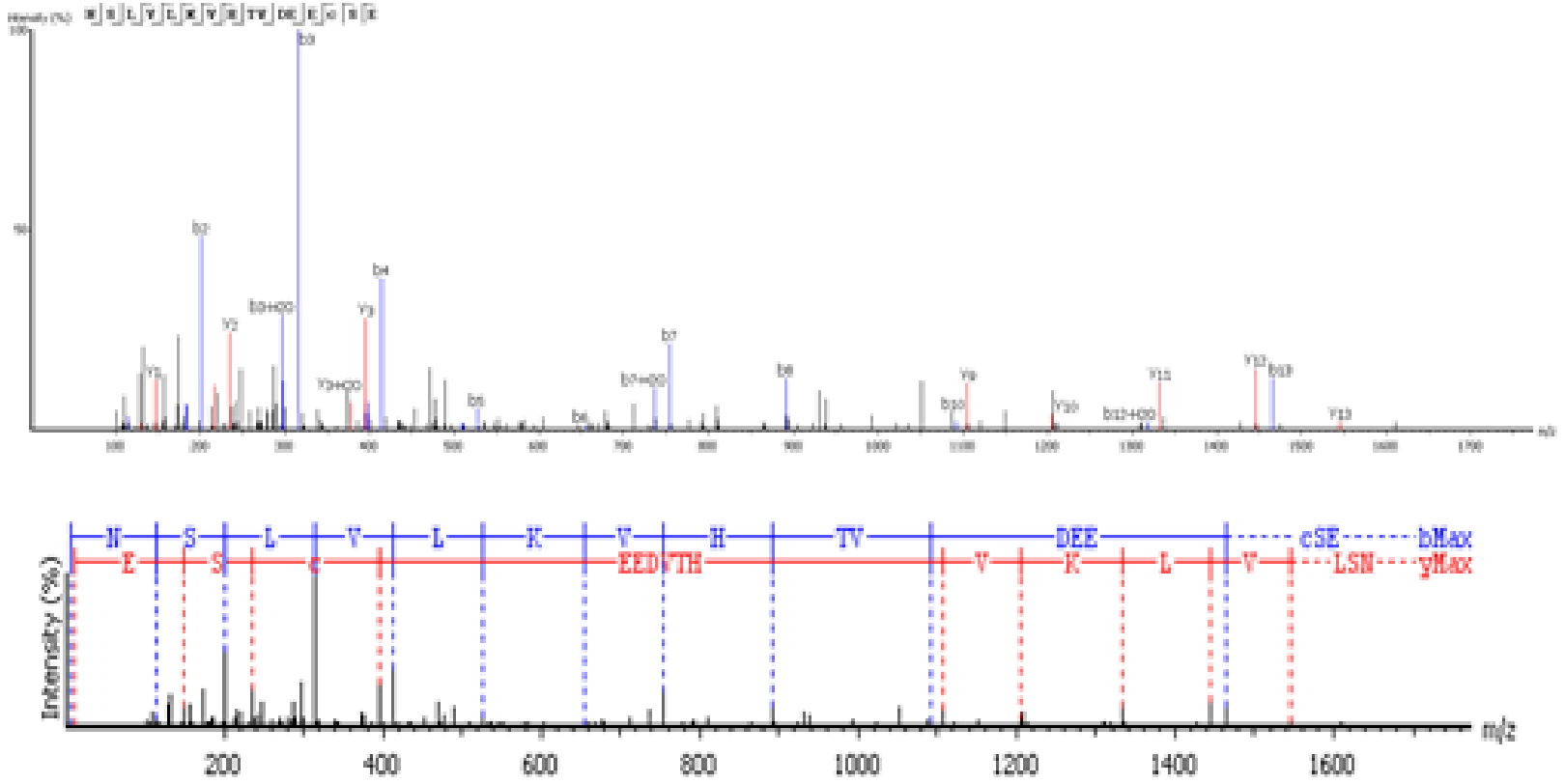
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 746 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



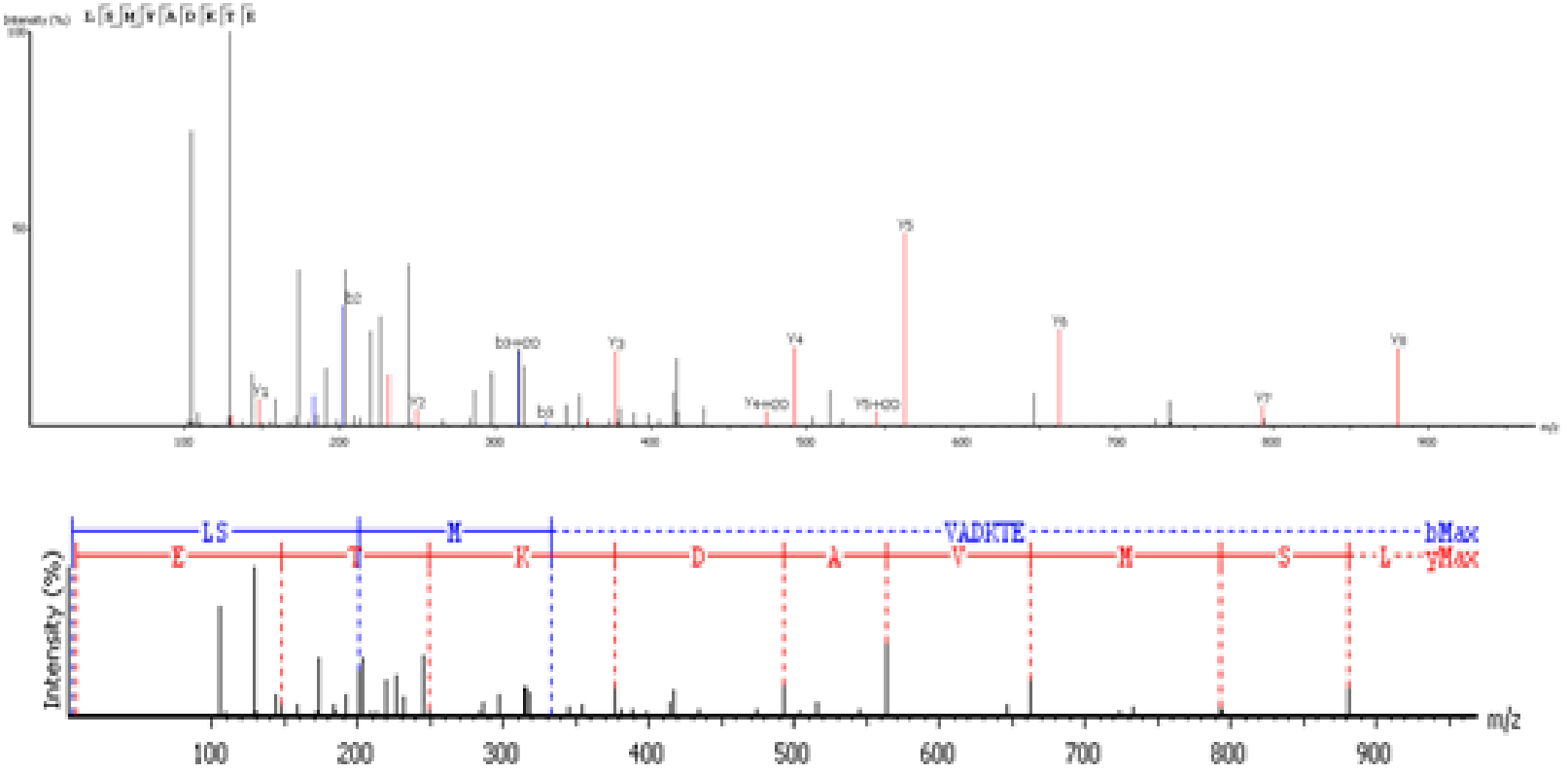
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1482 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



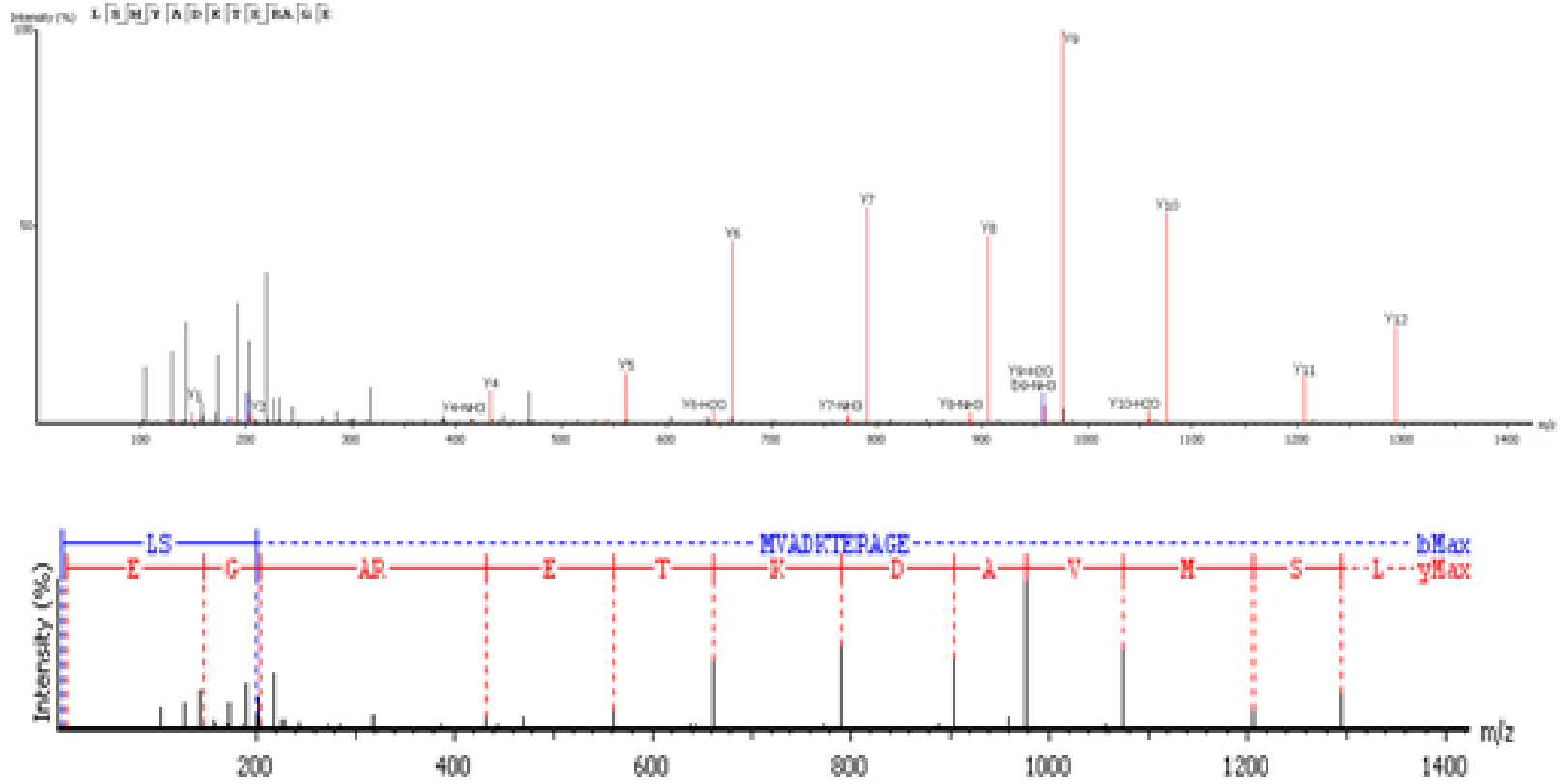
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1858 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



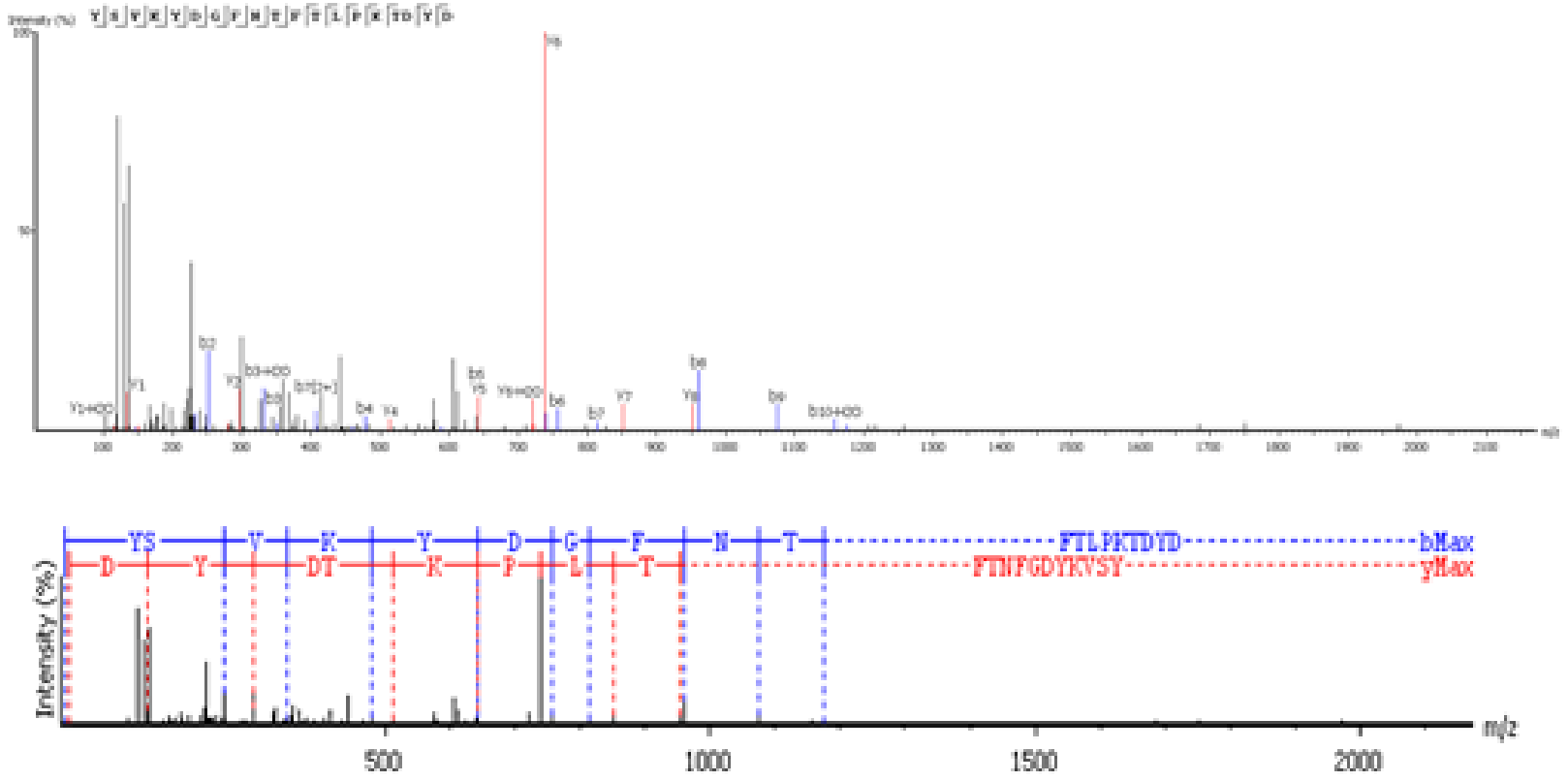
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 992 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



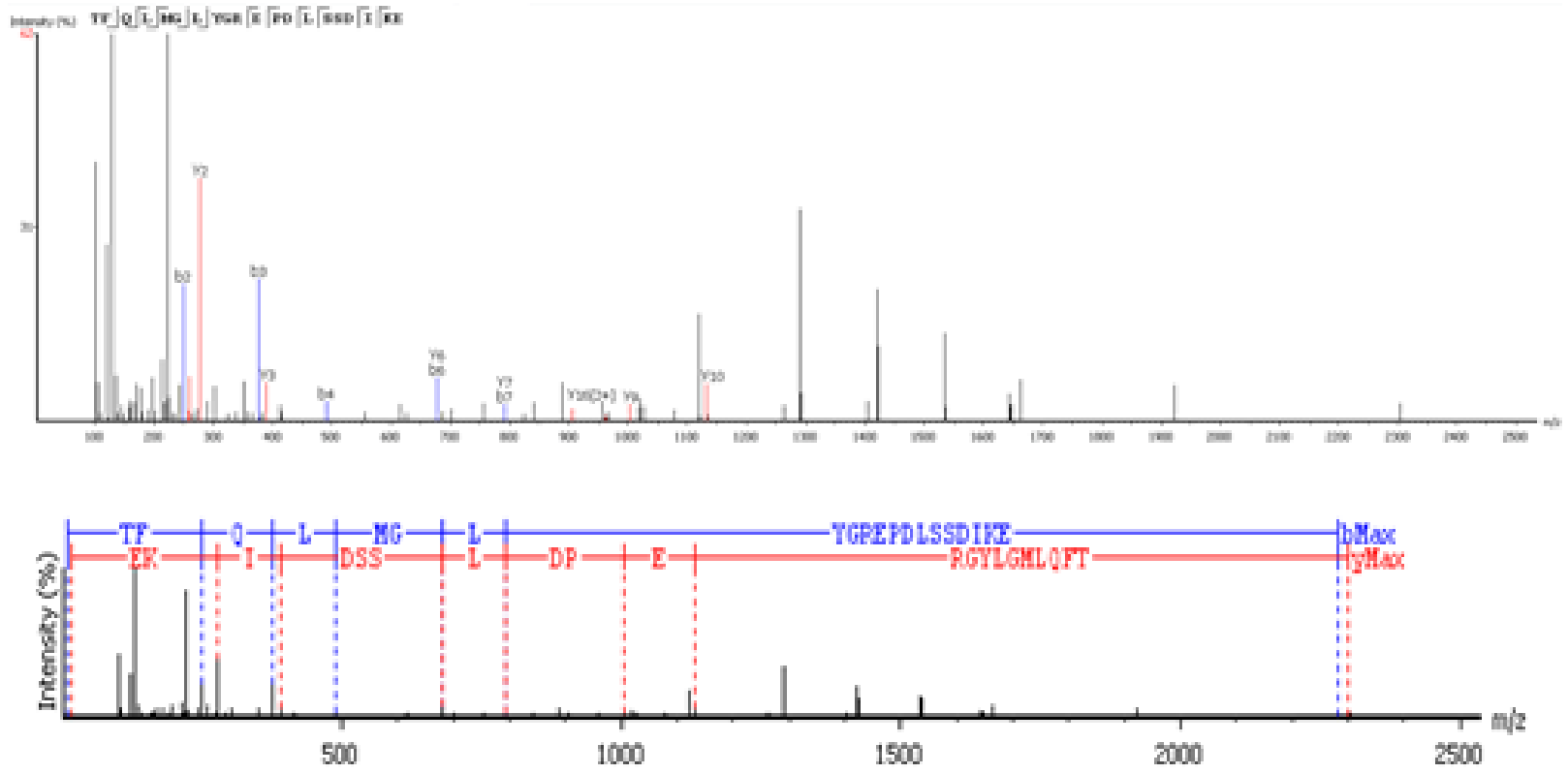
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1406 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



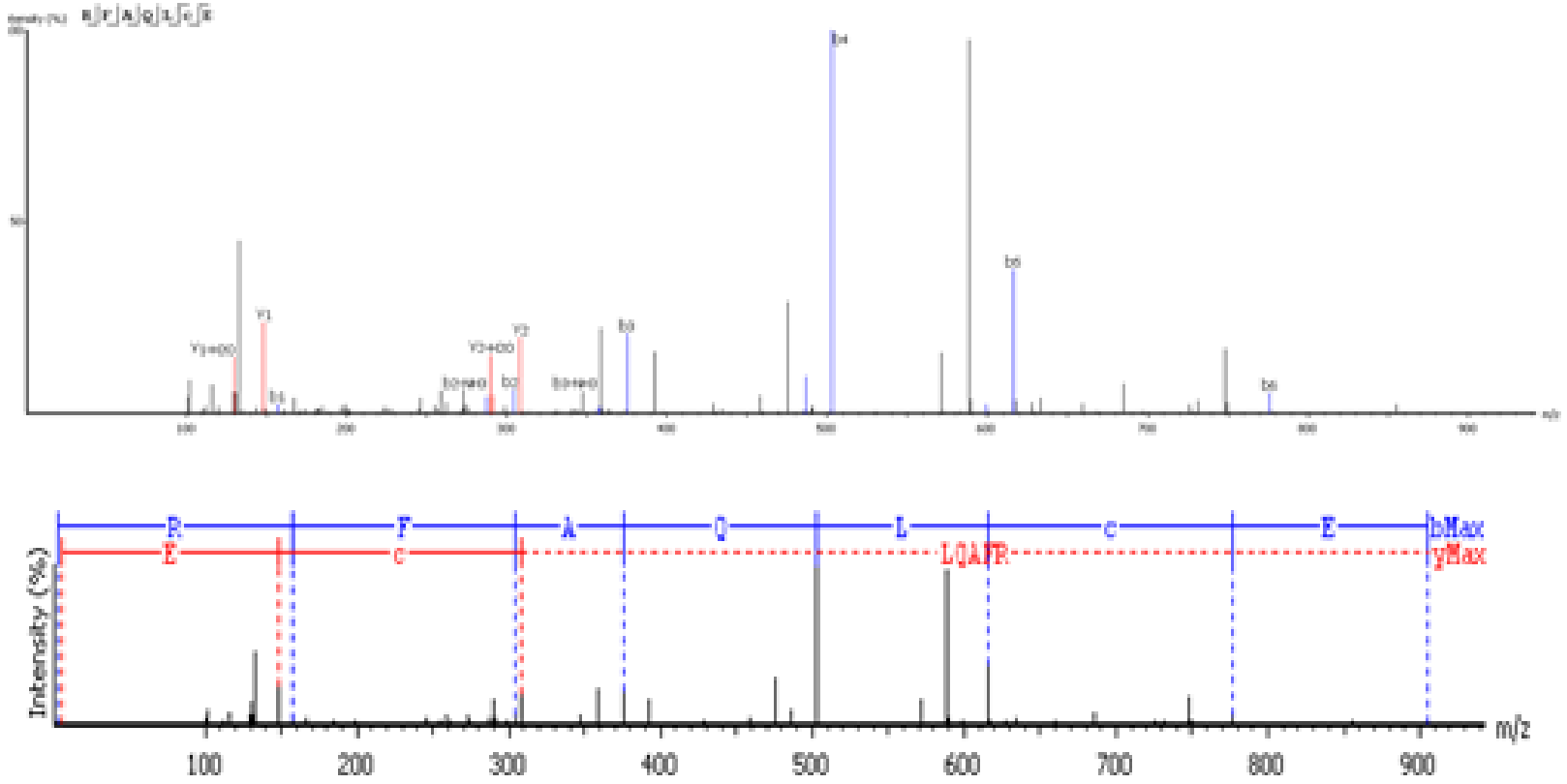
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 2274 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



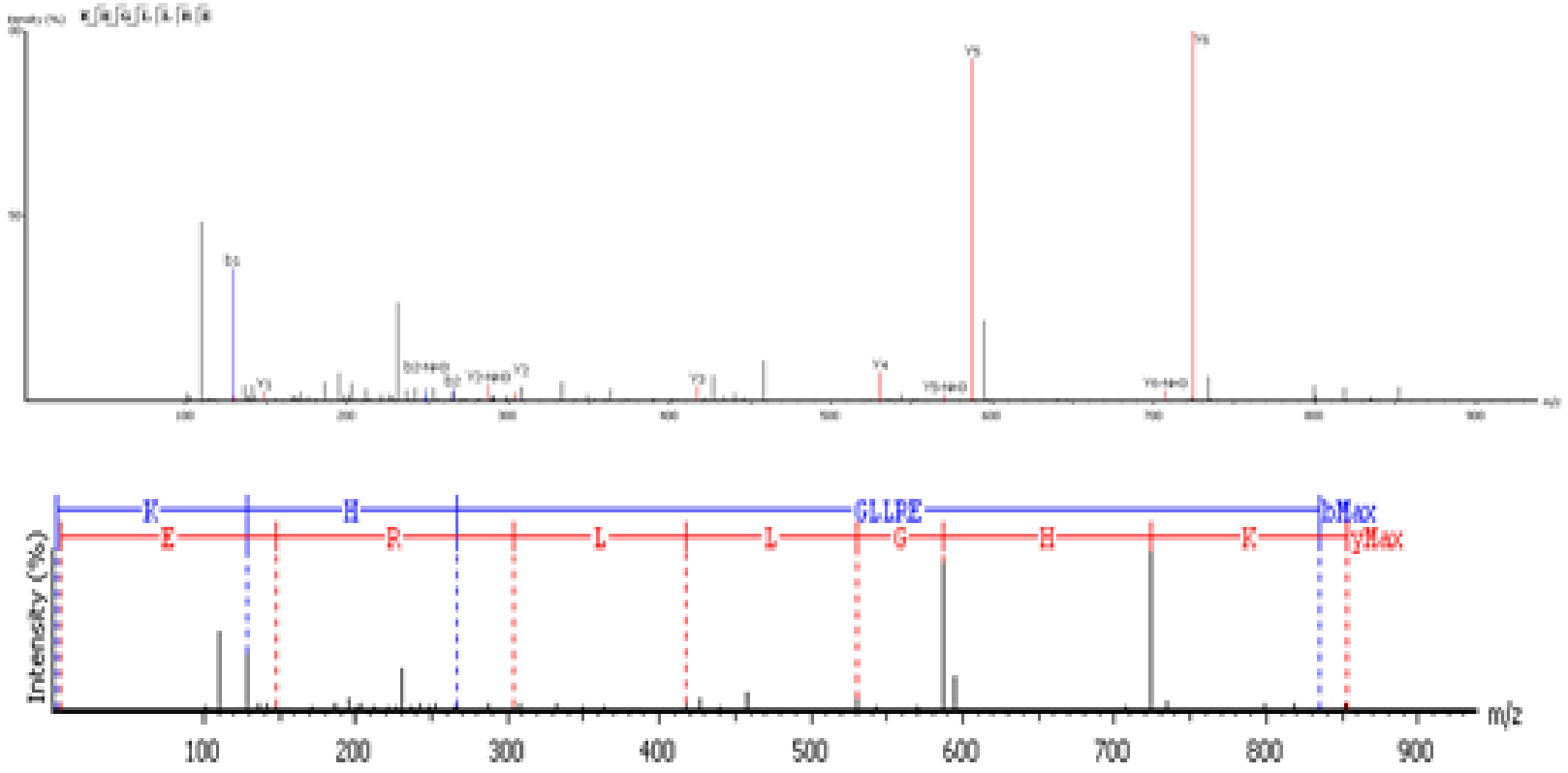
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 2298 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



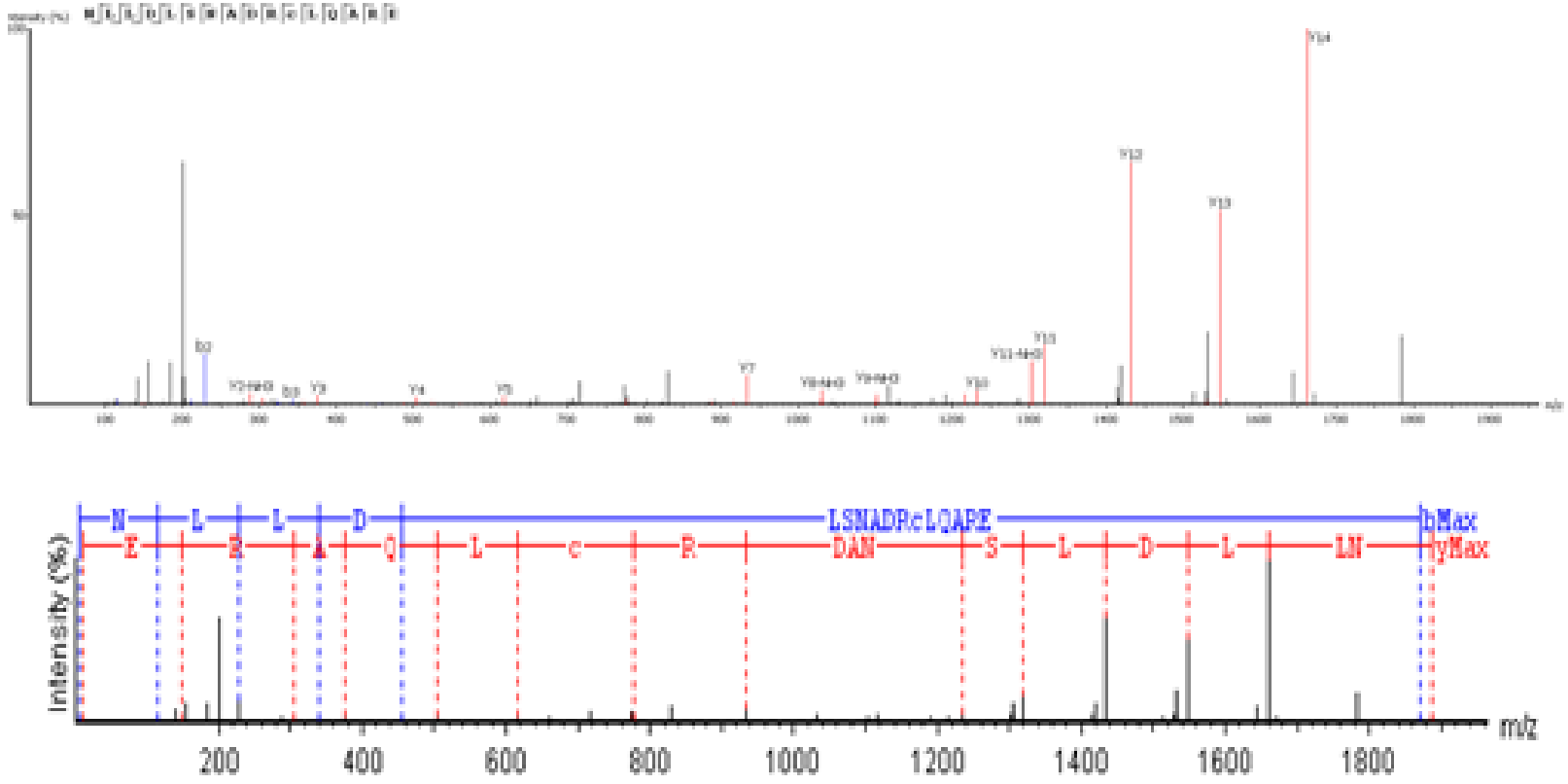
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 922 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



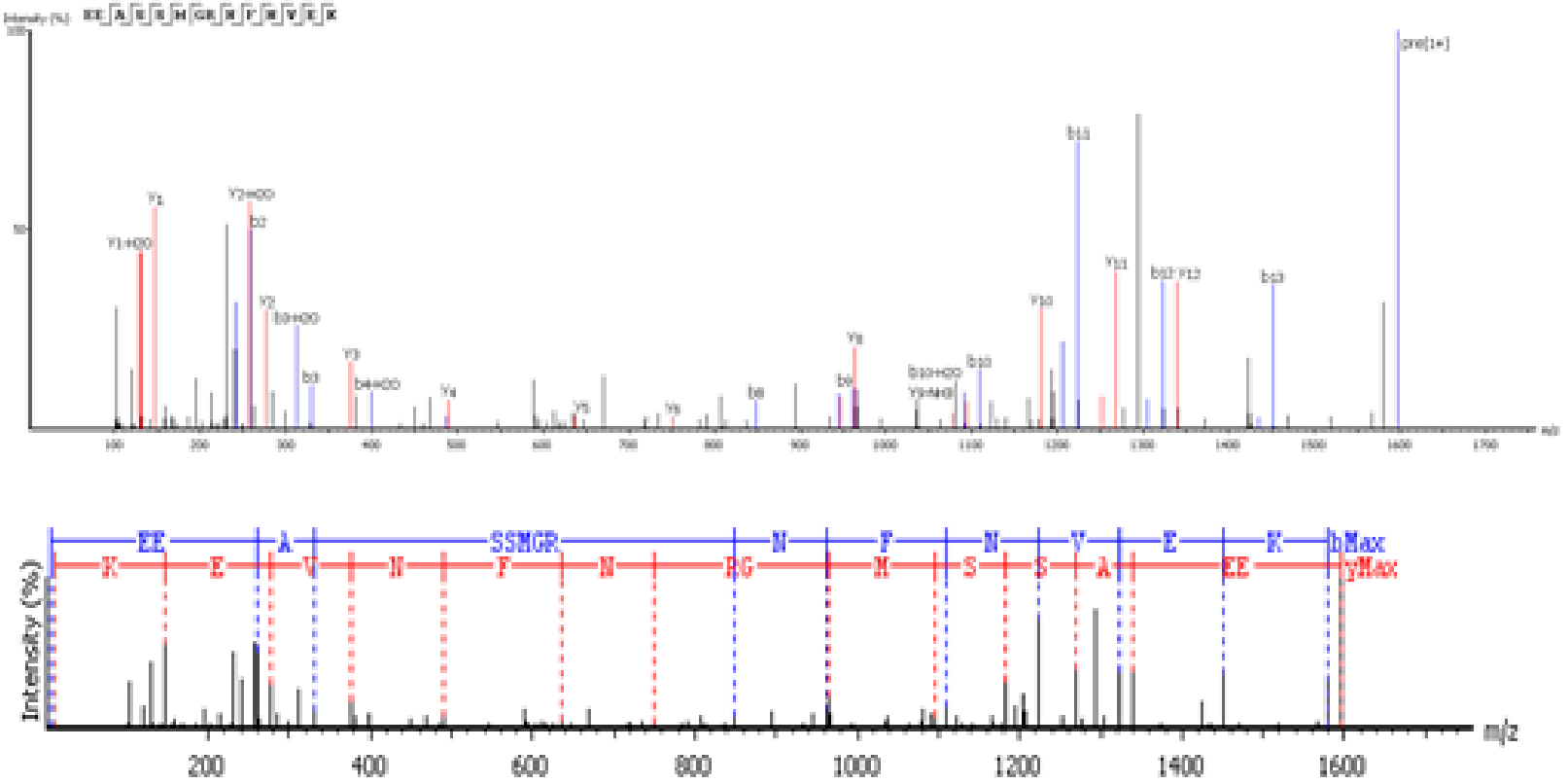
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 851 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



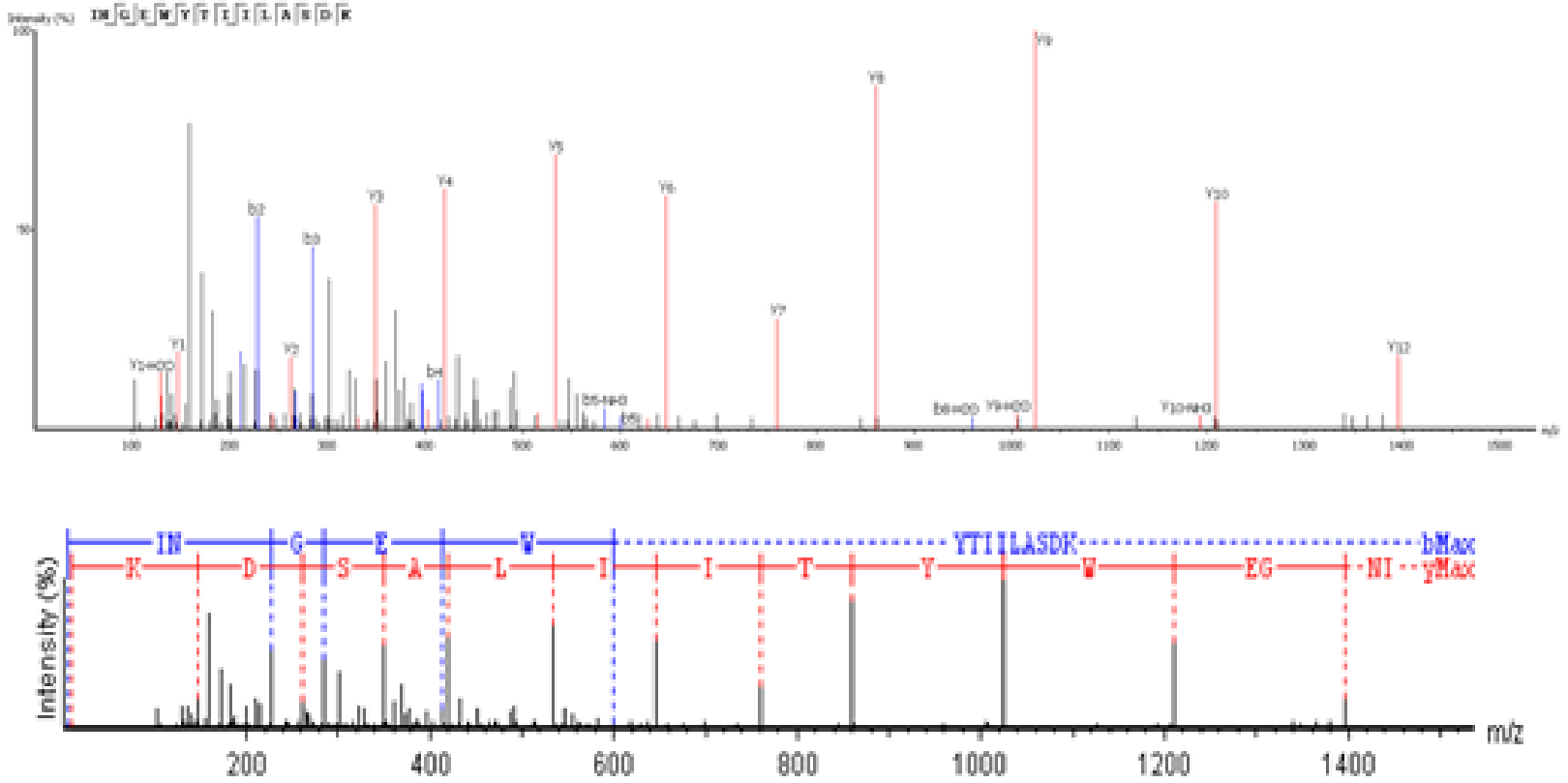
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1887 Da of 18585 Da MUP.

Fraction 2 collected from SAX chromatography was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



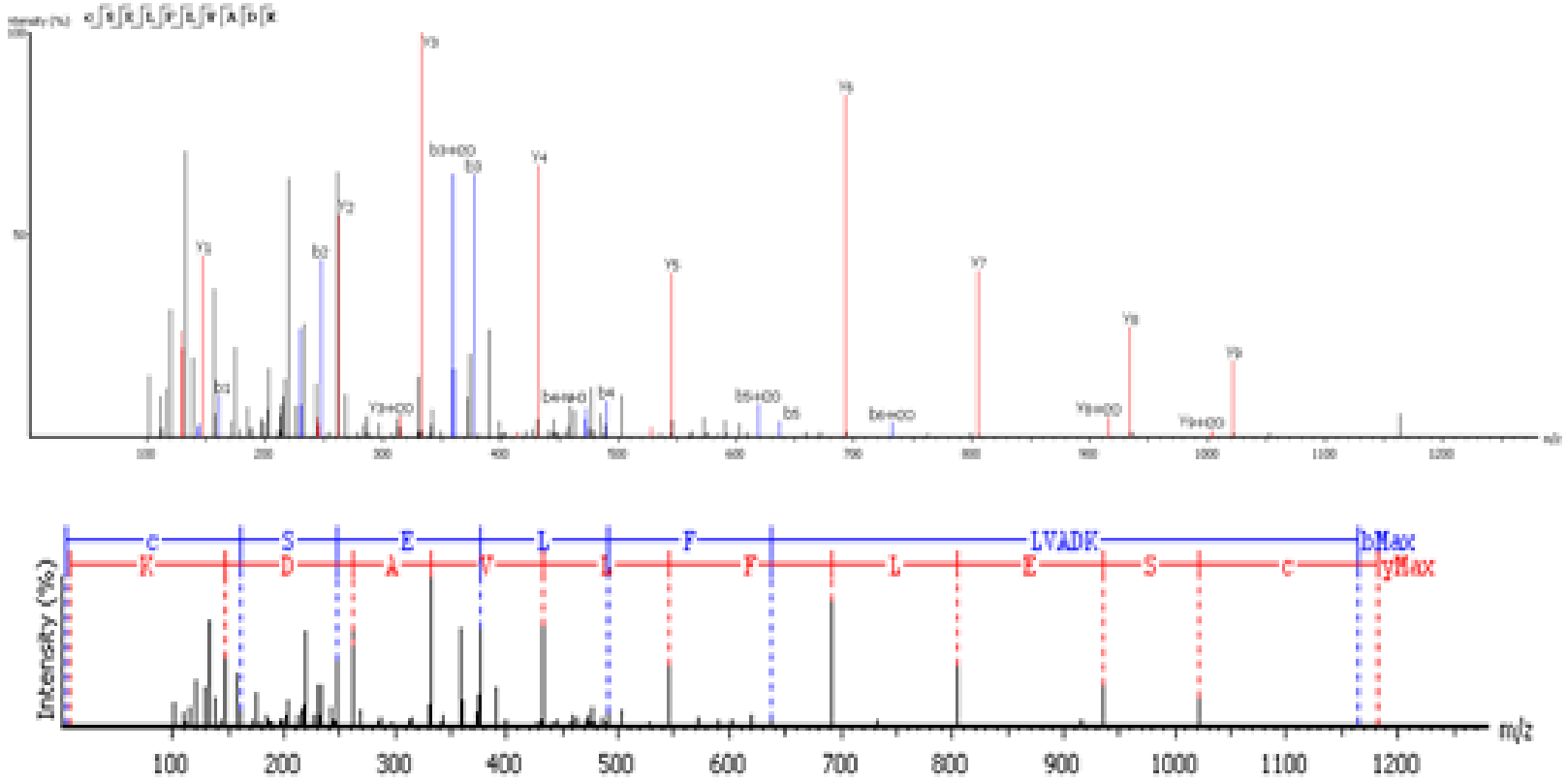
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1597 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



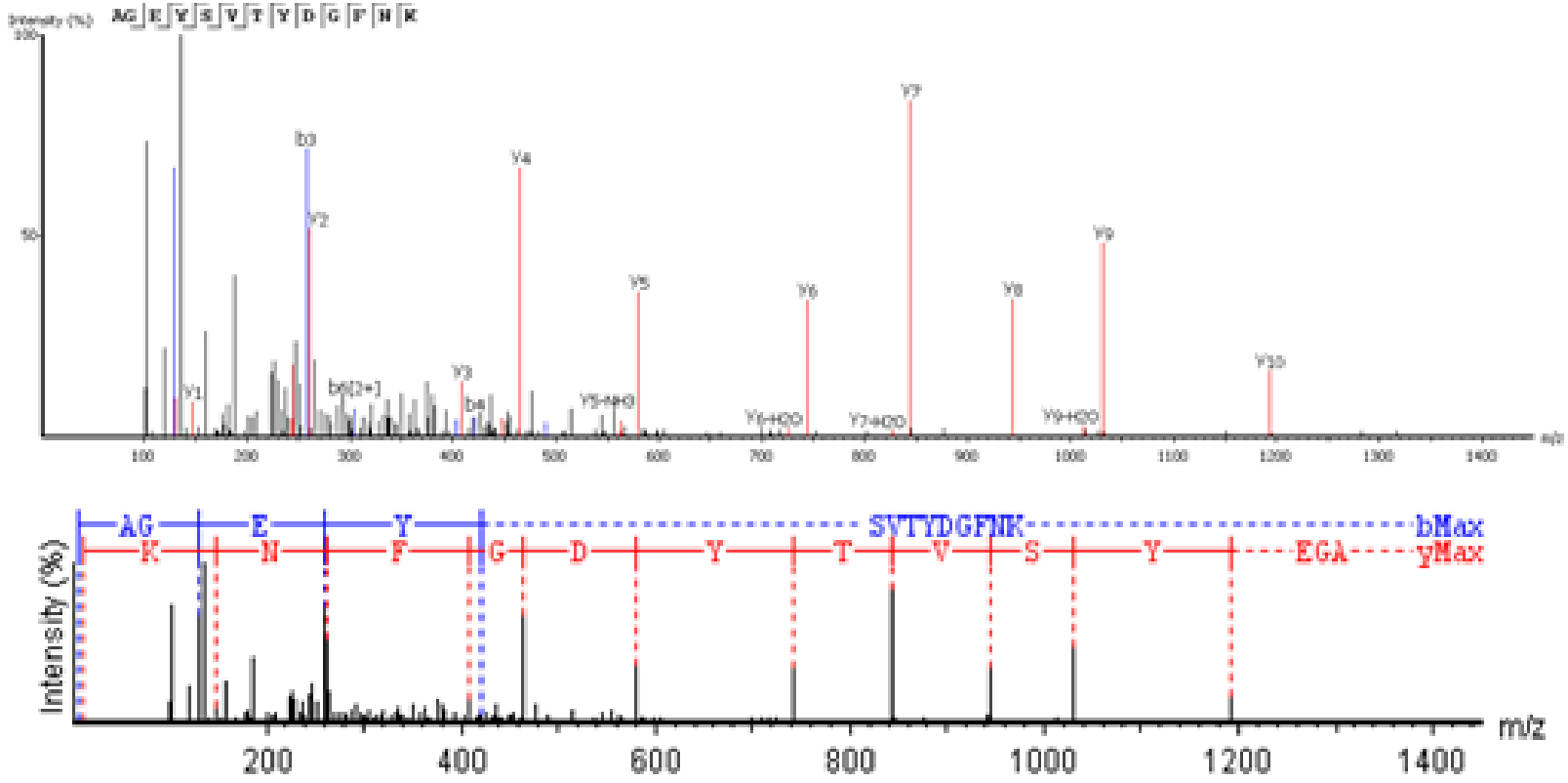
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1622 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



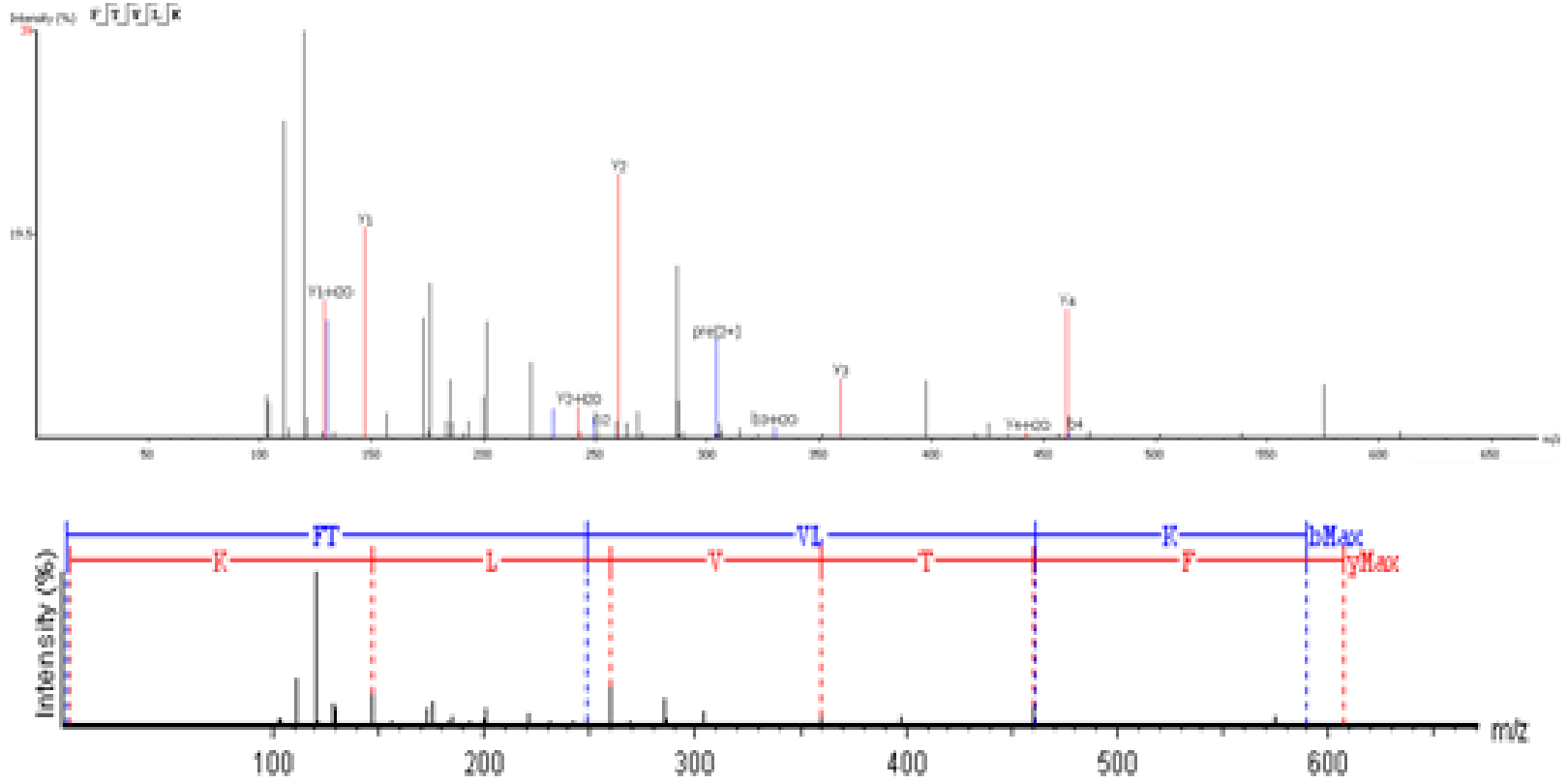
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1181 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



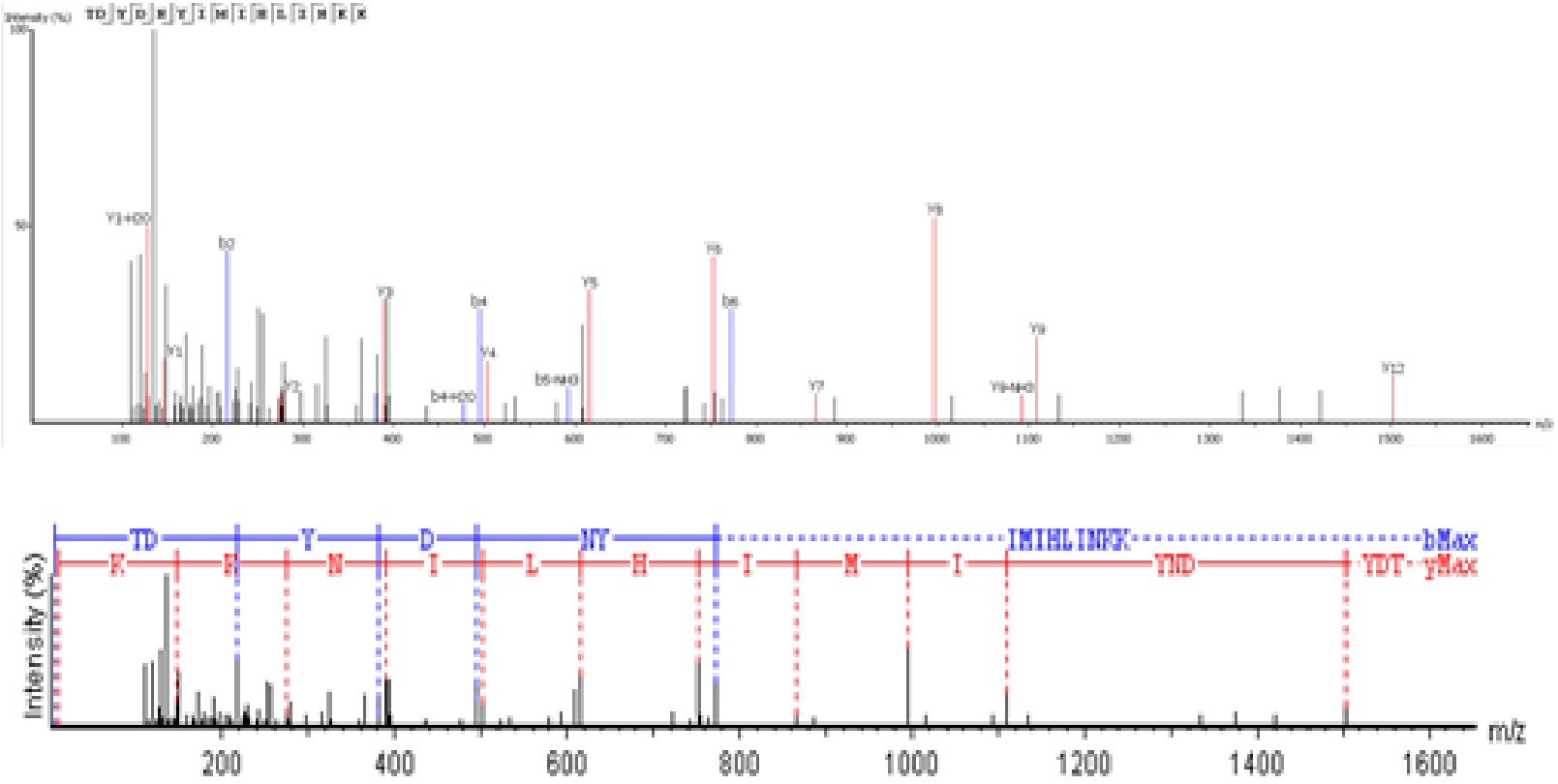
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1450 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



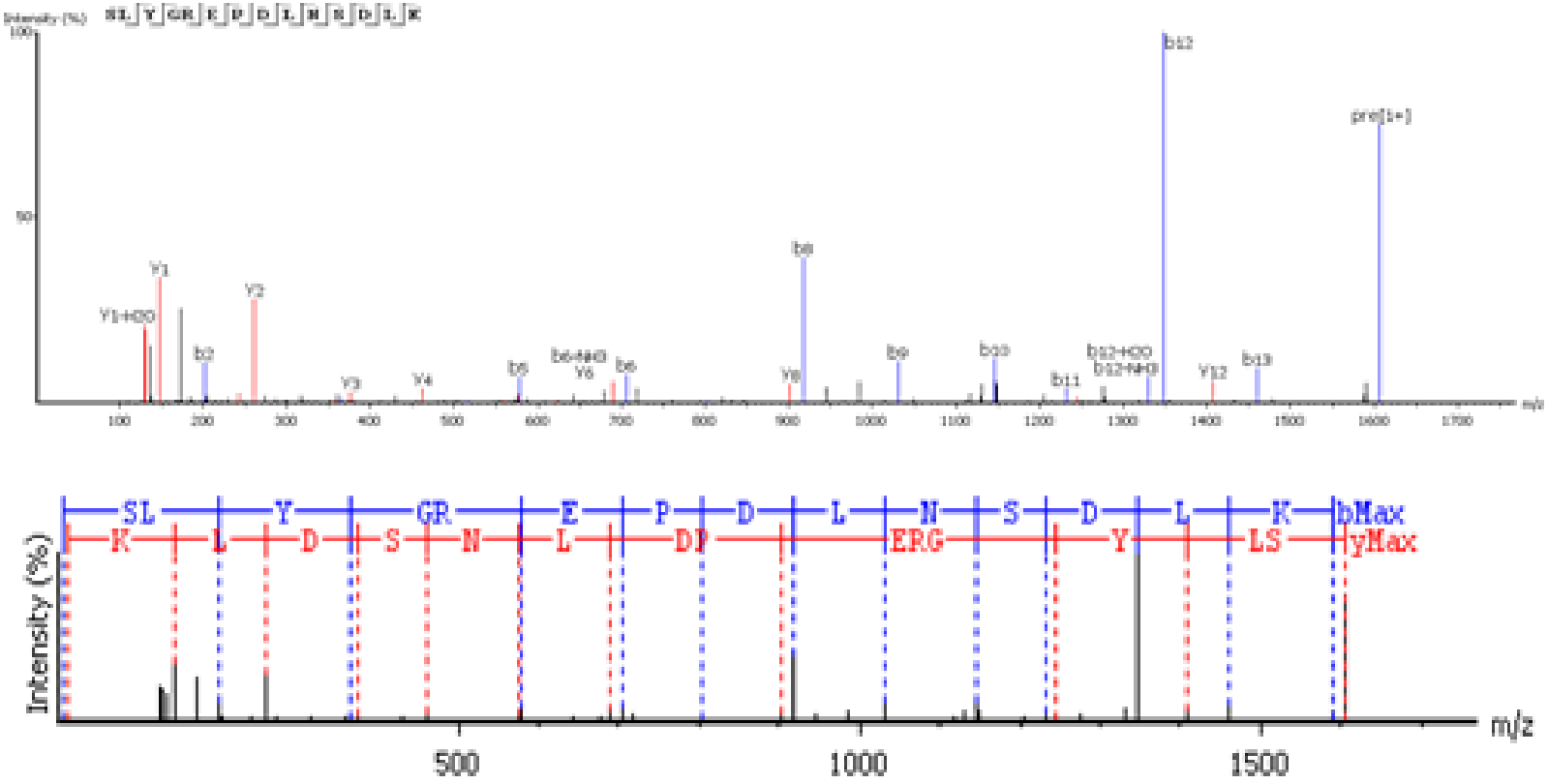
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 606 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



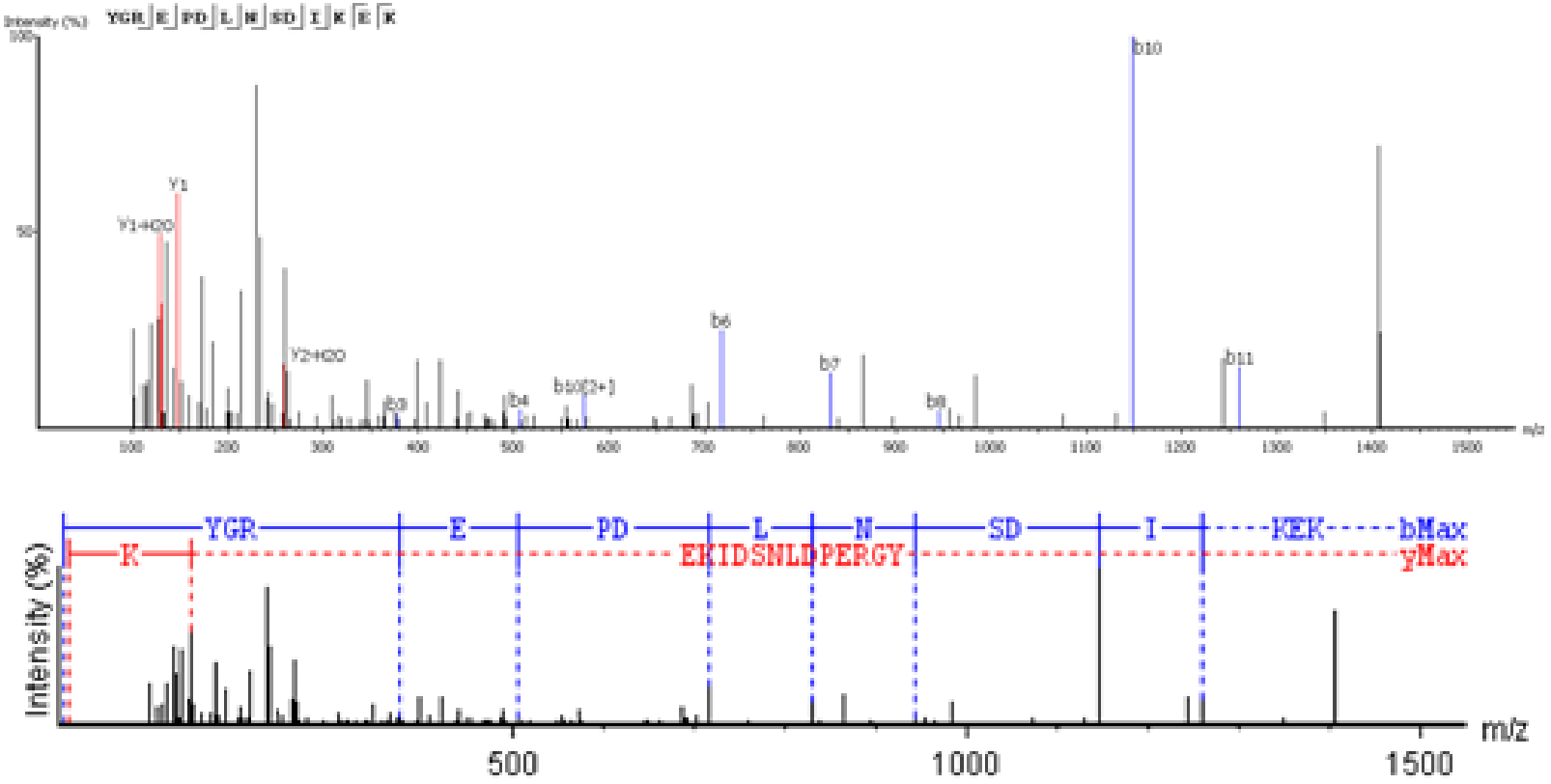
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1880 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



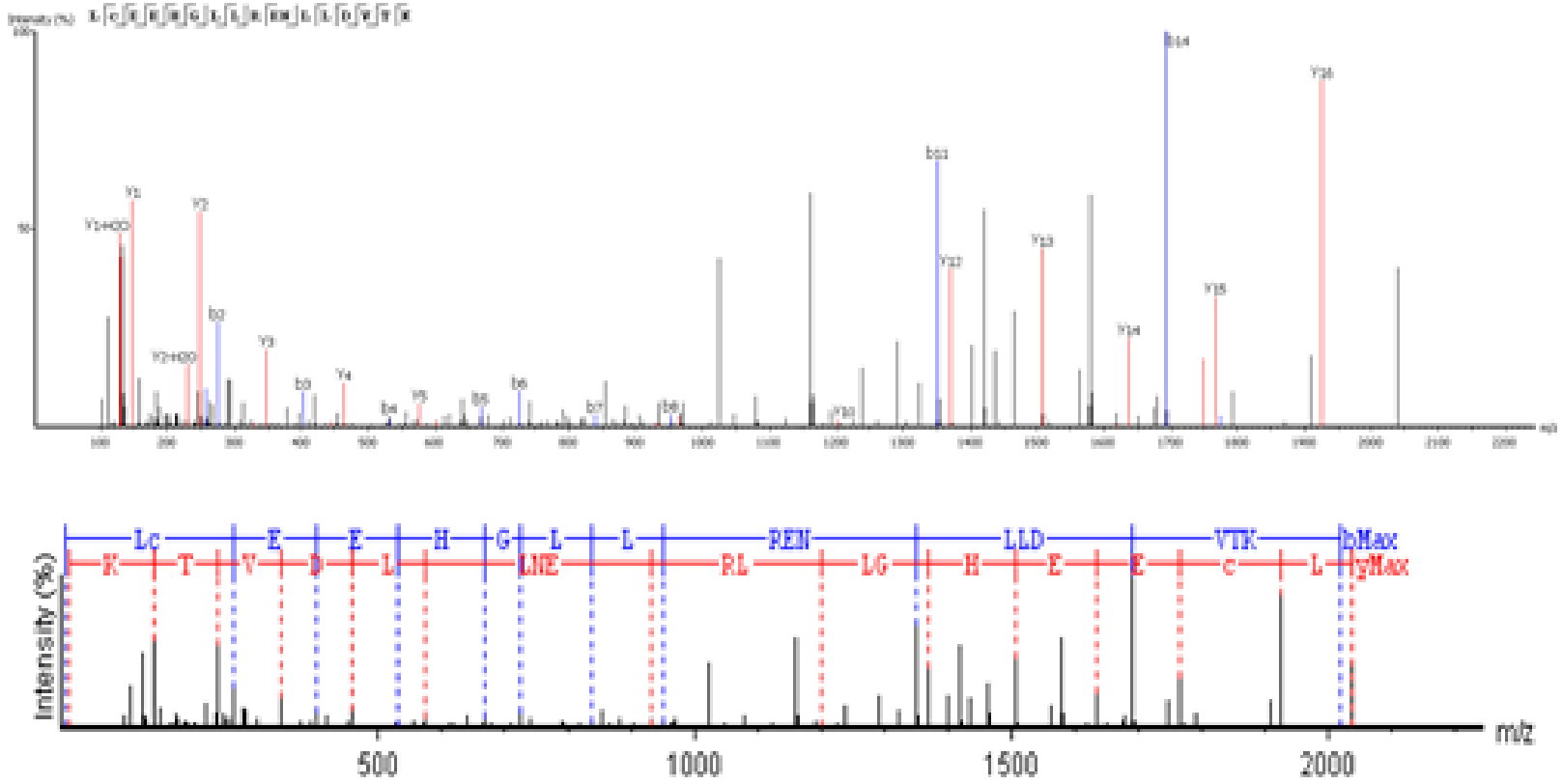
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1606 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



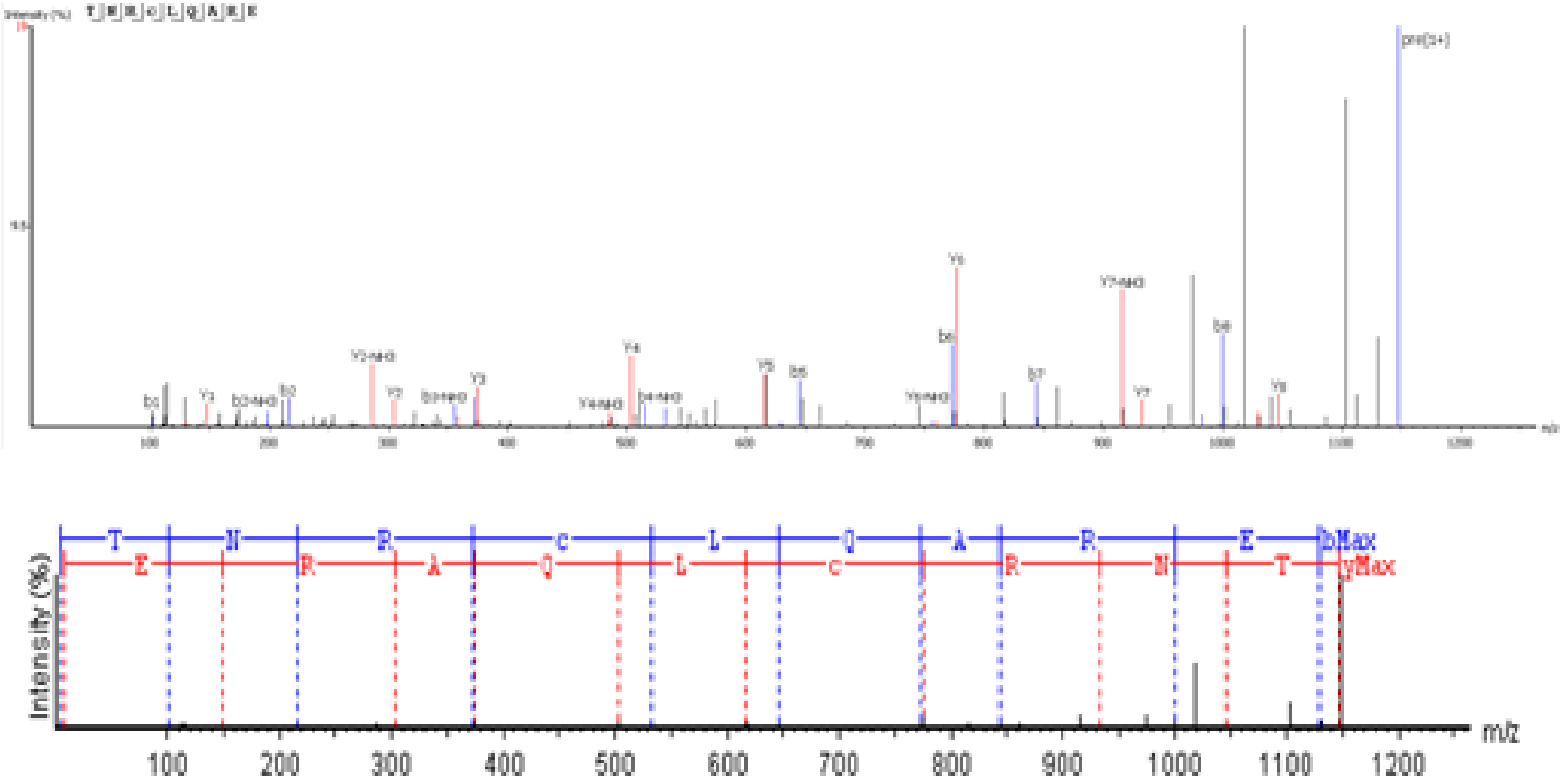
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1663 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



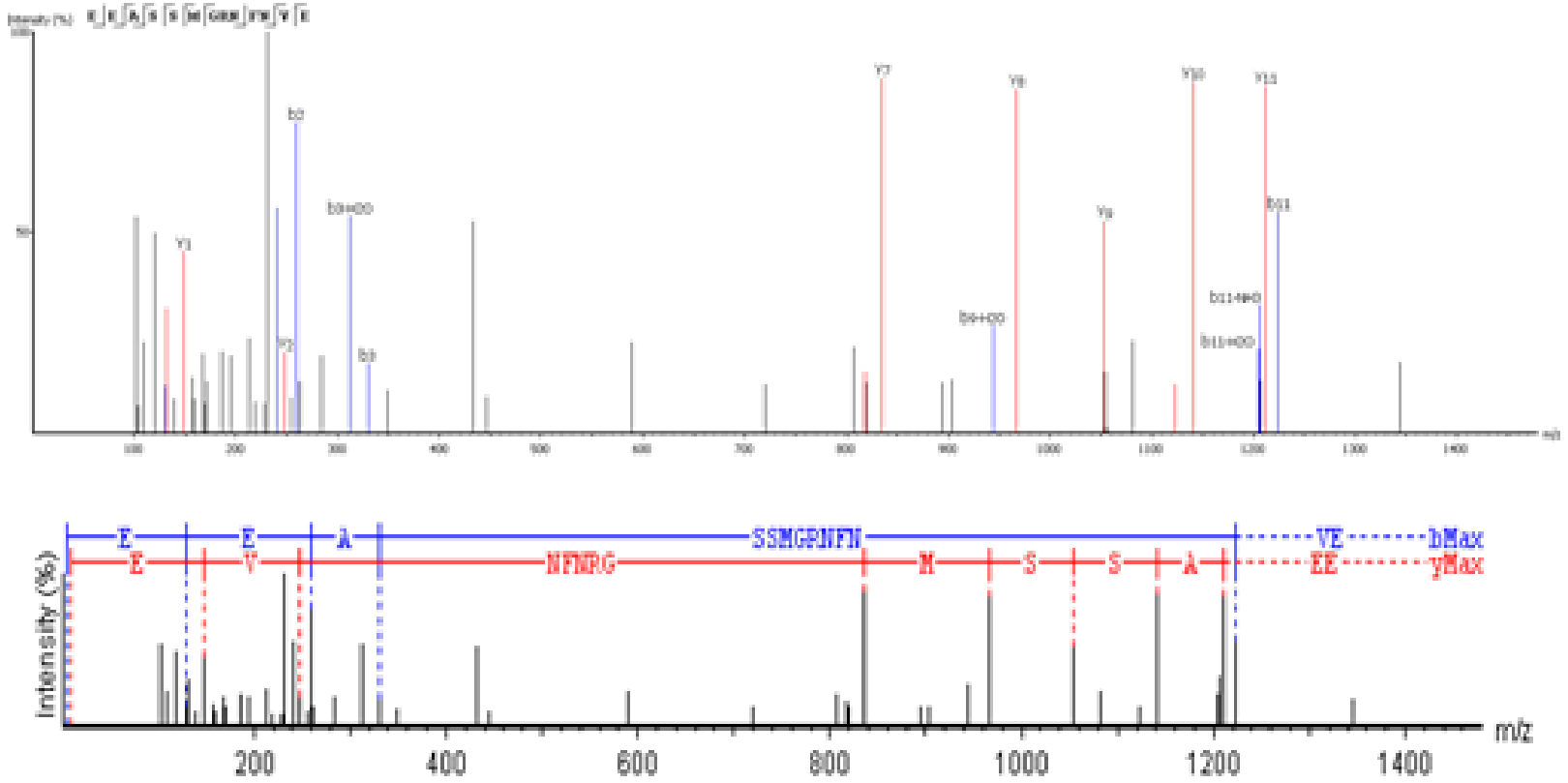
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 2038 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



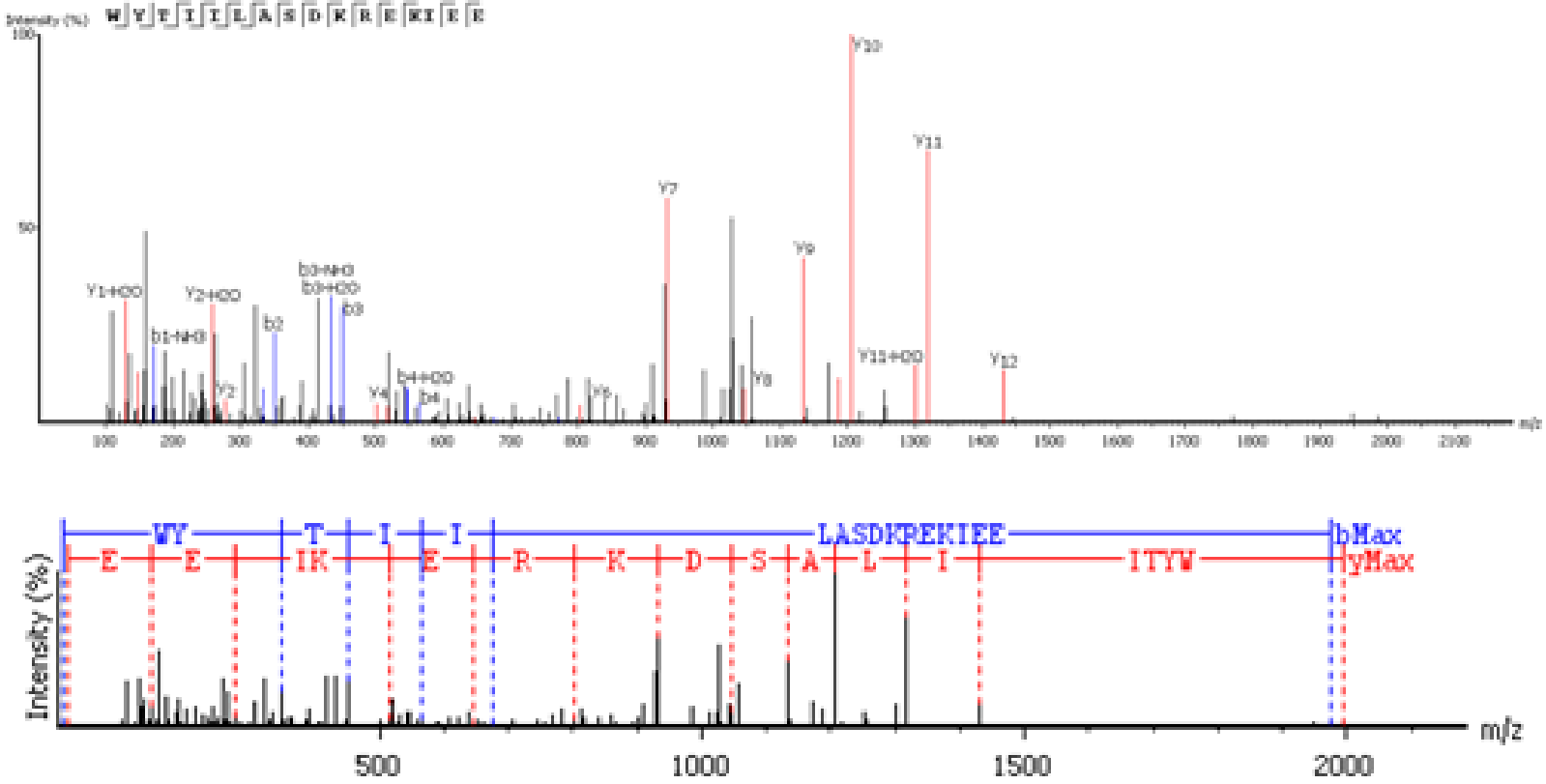
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Lys-C peptide 1147 Da of 18918 Da MUP.

Female urine was digested in-solution using Lys-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



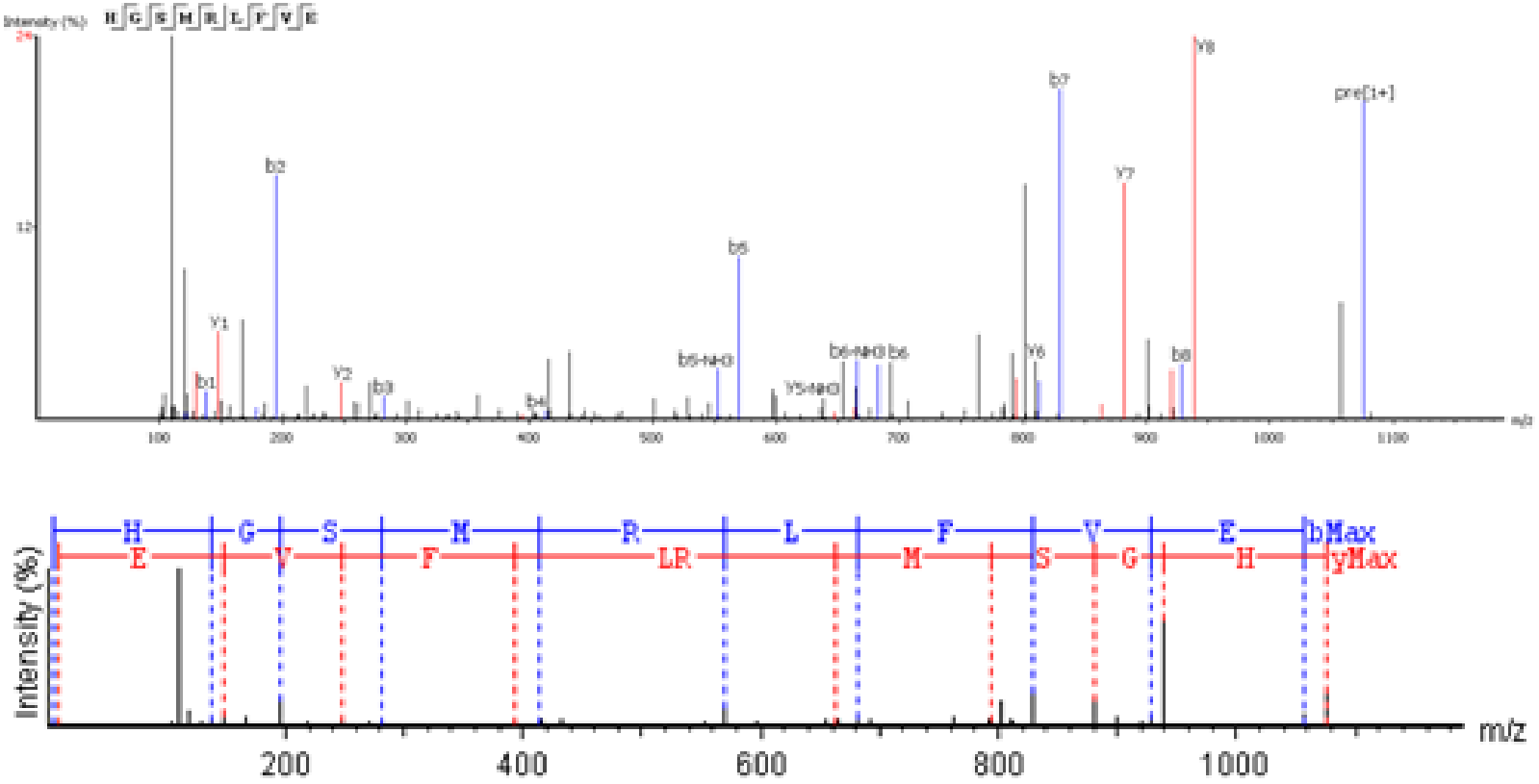
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1469 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



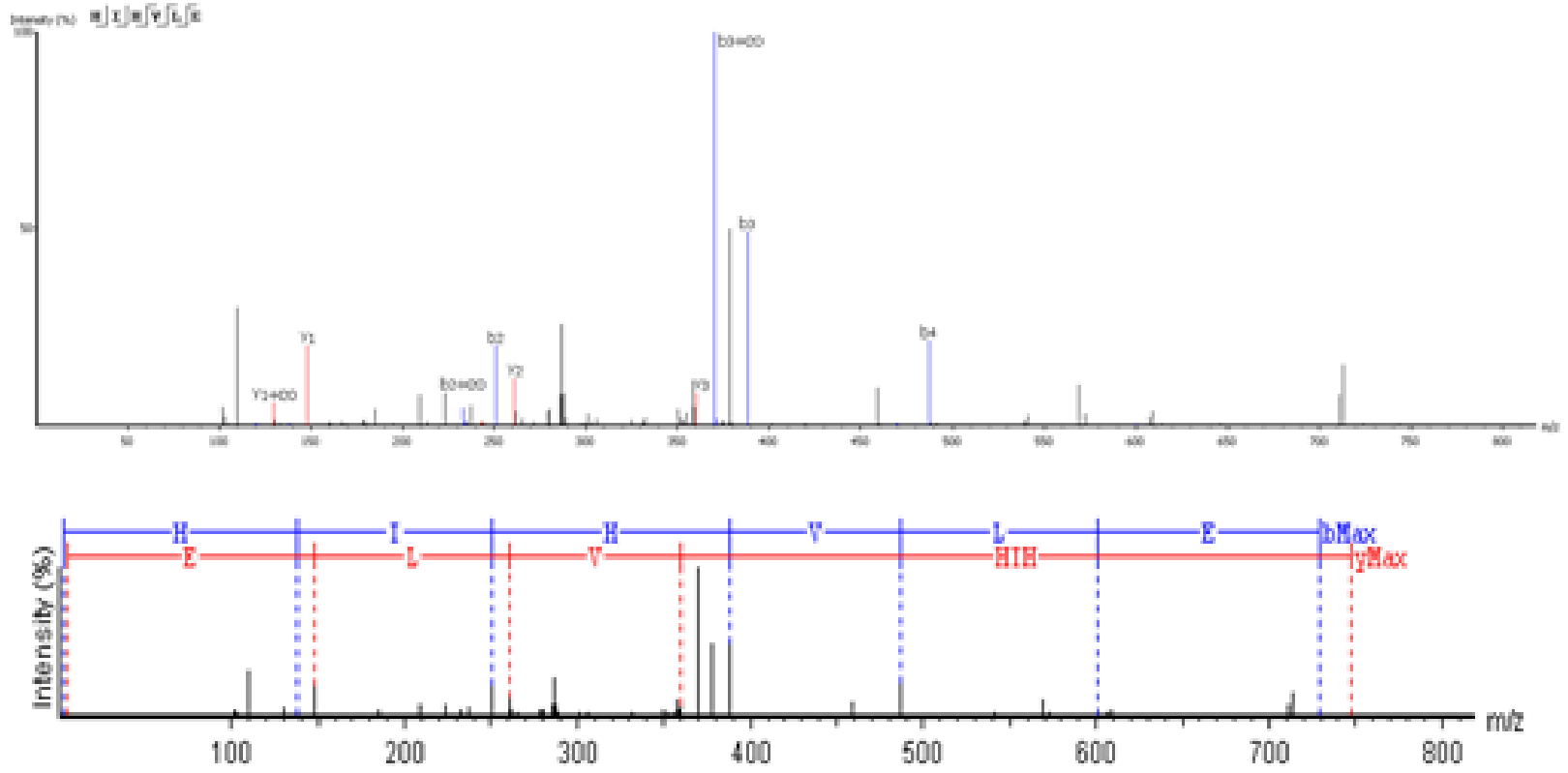
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1993 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



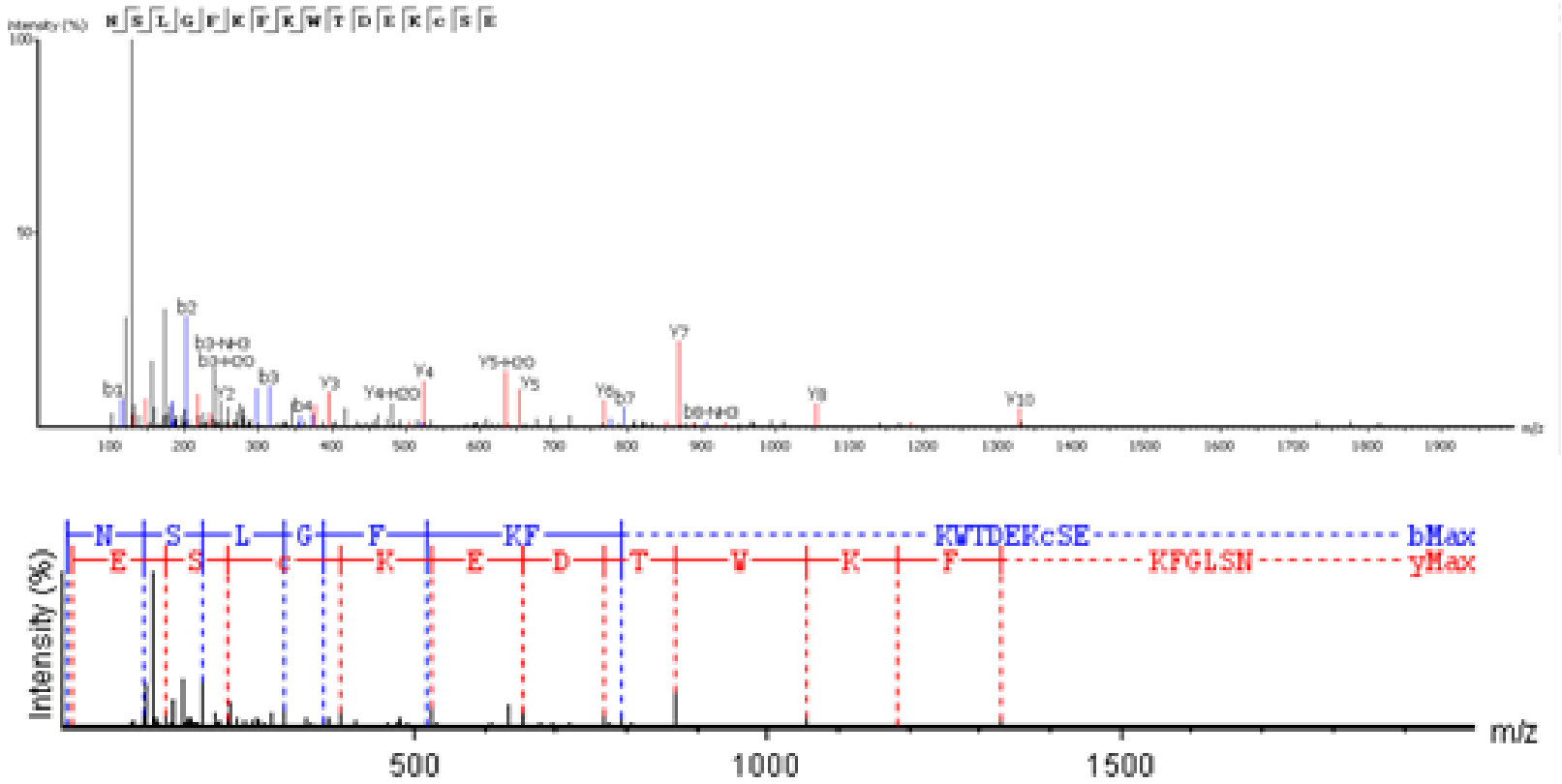
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1075 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



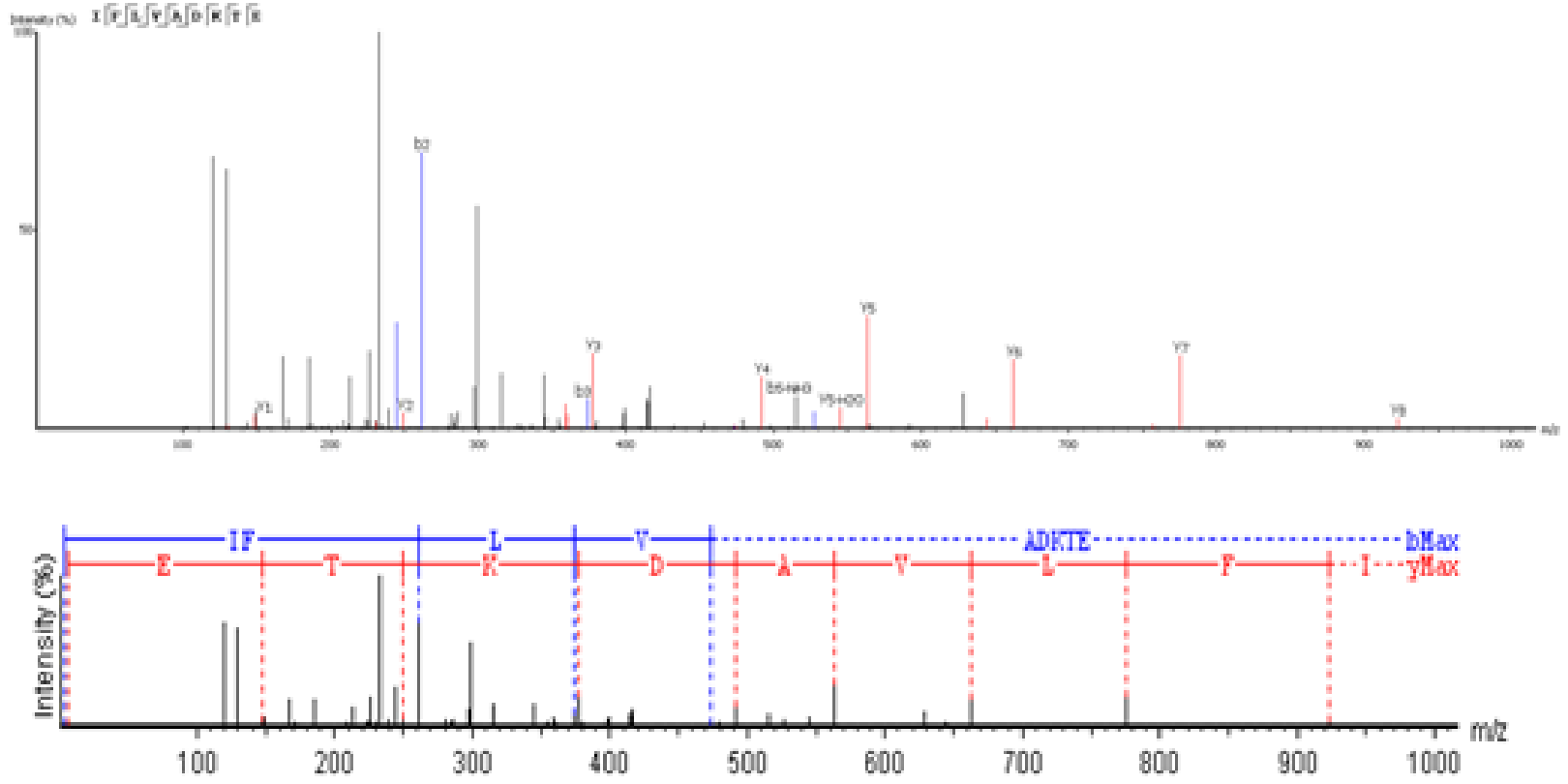
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 746 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



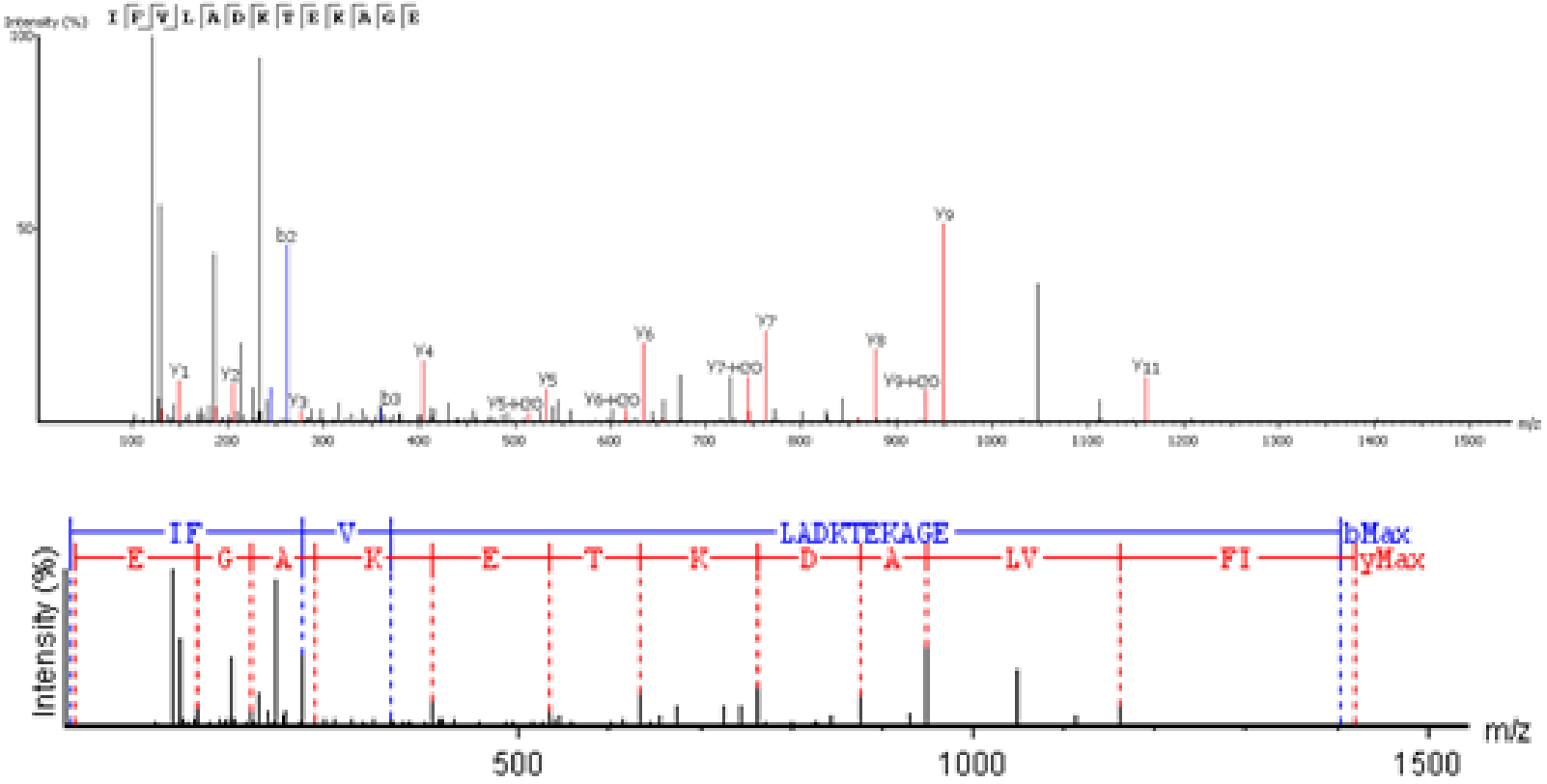
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1975 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



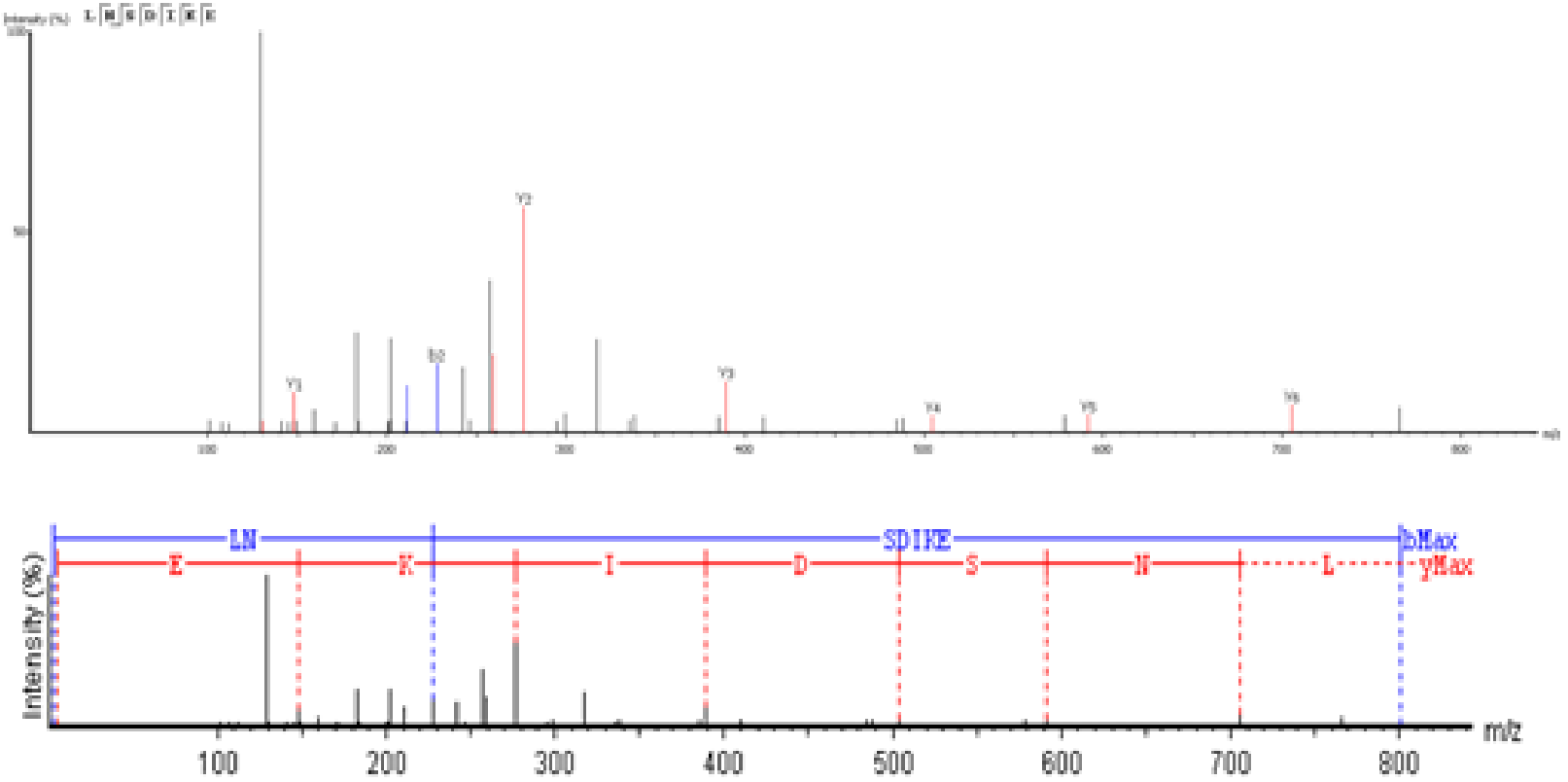
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1035 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



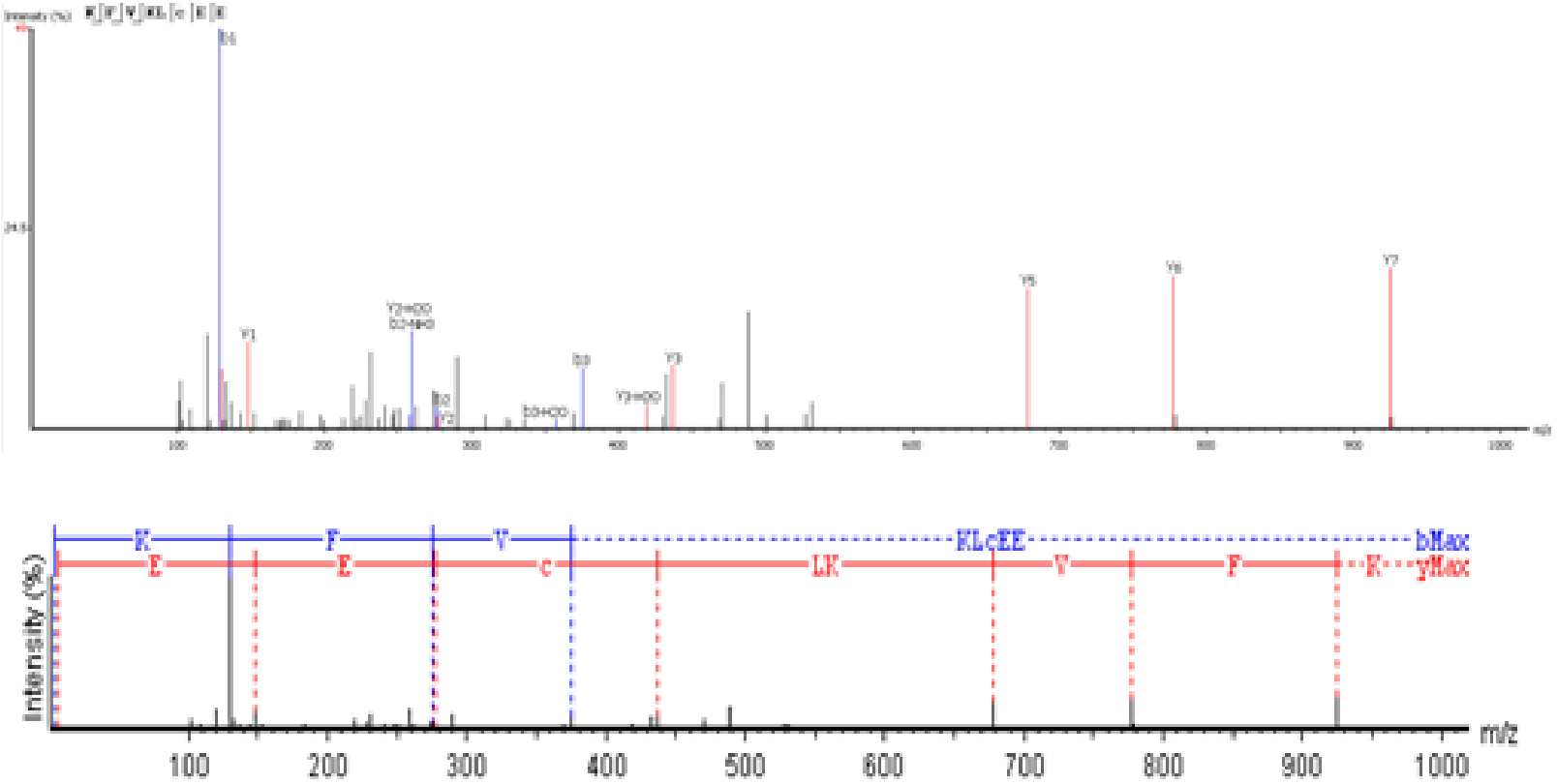
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1420 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



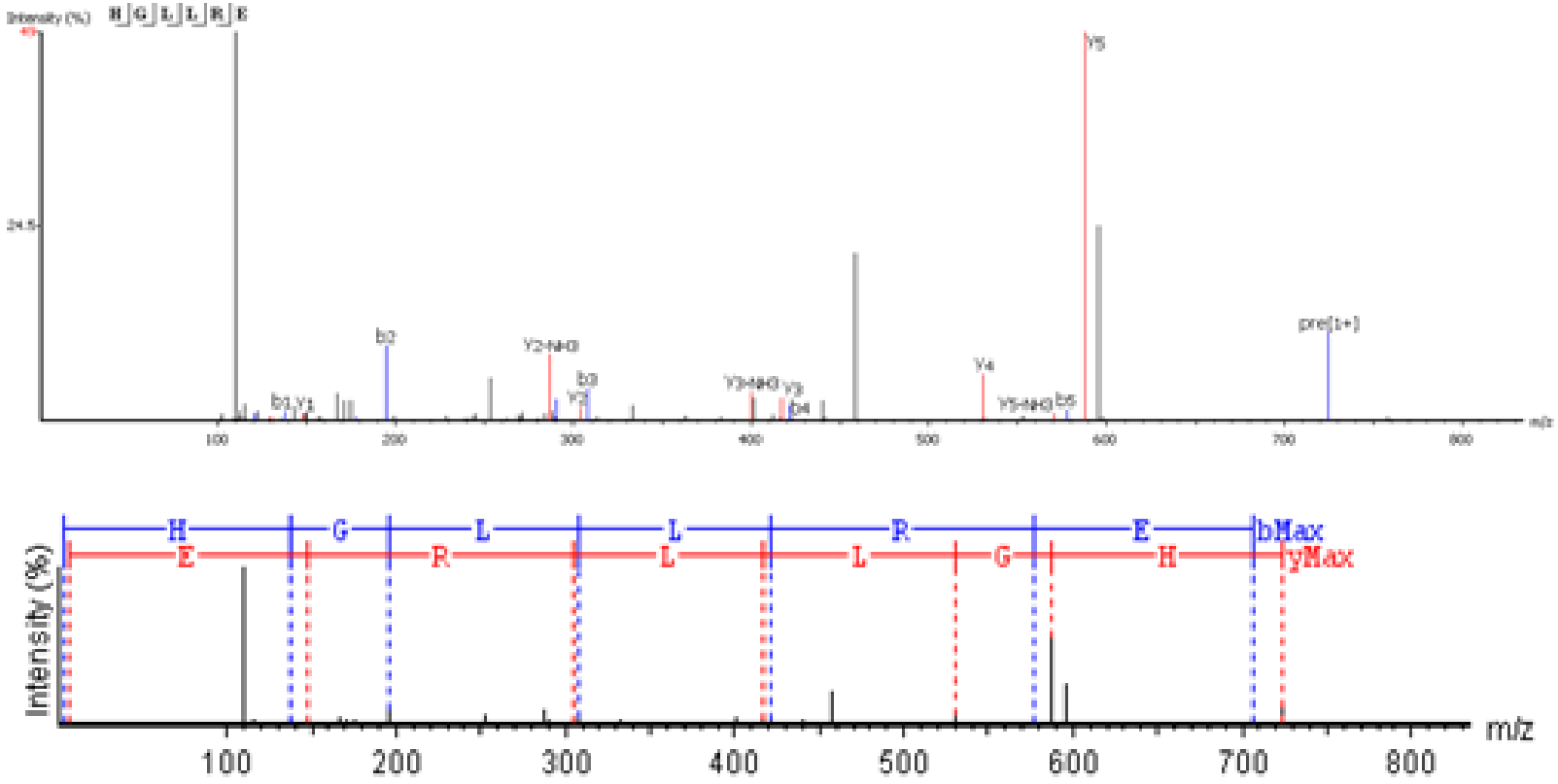
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 817 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



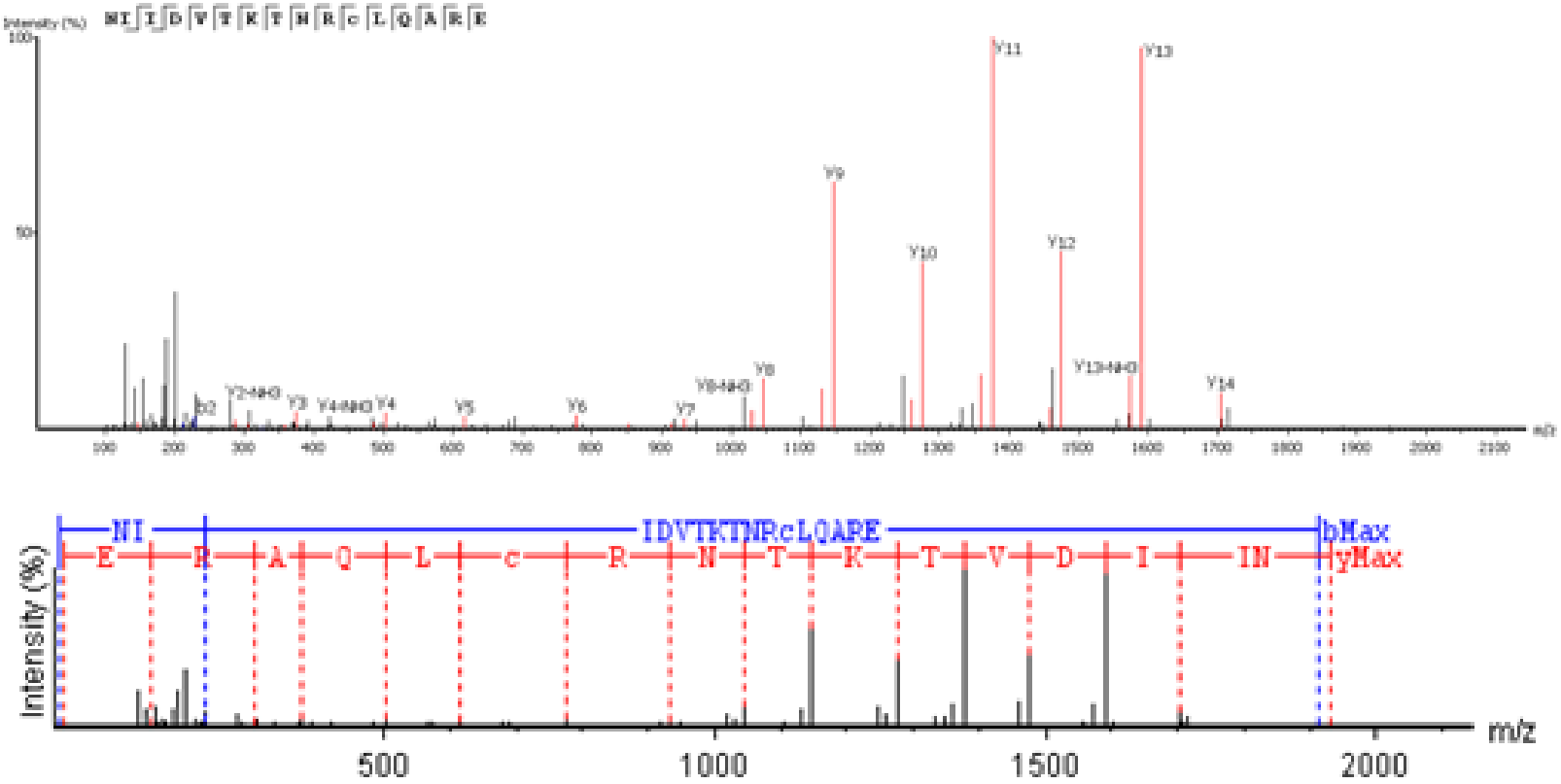
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1052 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



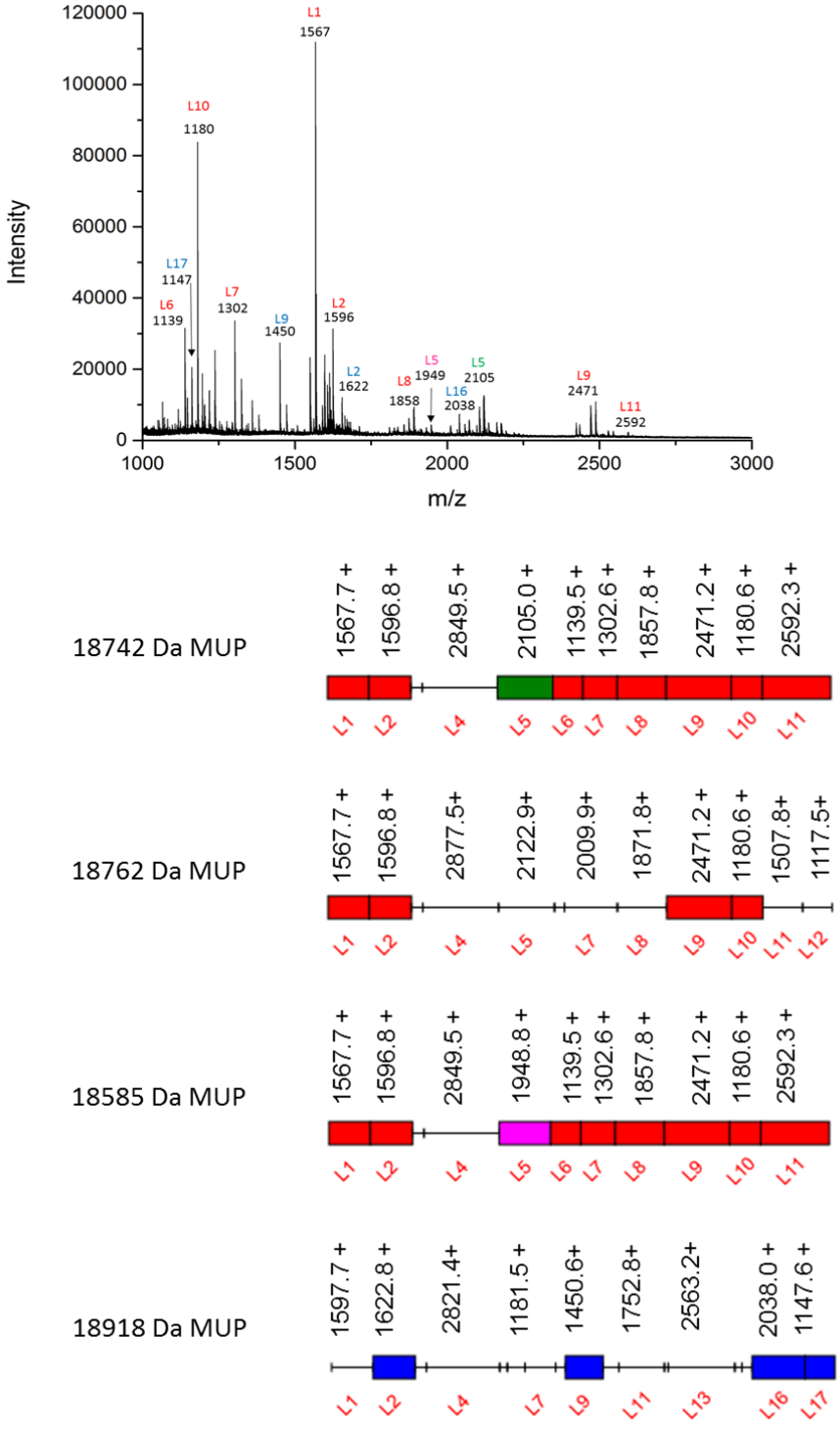
Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 723 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada). PEAKS software assigns leucine for all leucine and isoleucine residues, as the LC-MS system cannot distinguish between these isobaric residues.



Supplementary material A: *De novo* sequencing using PEAKS software of the processed MS/MS spectra of Glu-C peptide 1930 Da of 18918 Da MUP.

Female urine was digested in-solution using Glu-C as described in Chapter 2. Resulting peptides were analysed by LC-MS/MS using the Thermo QExactive as described in Chapter 2. Spectra were acquired between 300-2000m/z. Raw data was processed and peptide sequence was obtained using PEAKS 6 ®software (Bioinformatics Solutions Inc, Canada).



Supplementary material A: *Mus spicilegus* MUP peptide mass fingerprinting of male *Mus spicilegus* urine sample

The MALDI-ToF-MS spectrum generated from a Lys-C in-gel digestion of SDS-PAGE gel piecesfrom SDS-PAGE analysis of male urine (top band) was searched against a custom made *Mus spicilegus* MUPs database using Mascot. The digested protein band from male urine showed peptides that matched peptides from all *Mus spicilegus* MUPs, the peptide maps of which are shown below the spectrum. In the peptide maps, the coloured peptides are those masses observed in the spectrum. The red peptides are those with masses shared between two or more of the *Mus spicilegus* MUPs, and peptide masses unique to each MUP are coloured differently to one another – peptide masses unique to 18742 Da are coloured green; to 18762 Da, purple; to 18585 Da, pink; to 18918 Da; blue.

		Behaviour						
Interaction	Male	Sniffing	Approach	Retreat	Chasing	Escaping	Aggression	Avoidance
Day 1	A	✓	✓			✓		
	B	✓	✓		✓		✓✓✓	
Day 2	A						✓✓✓✓	
	B						✓✓✓✓	
Day 3	A						✓✓✓✓	
	B						✓✓✓✓	

Supplementary material A: Recording of the behaviour observed in the first male/male pair.
During the controlled interactions, the behaviour by both mice was observed and recorded in the above table. Each mouse was identified by a fur clip on either their right shoulder or right rump. A tick represents that that particular behaviour was observed. More ticks indicates more persistent behaviour. Aggressive behaviour was stopped by splitting the mice up to avoid injuries being sustained.

		Behaviour						
Interaction	Male	Sniffing	Approach	Retreat	Chasing	Escaping	Aggression	Avoidance
Day 1	A	✓	✓✓		✓			
	B	✓		✓		✓		✓
Day 2	A	✓	✓✓✓					
	B	✓	✓	✓				
Day 3	A	✓	✓					
	B	✓	✓					

Supplementary material A: Recording of the behaviour observed in the second male/male pair.
During the controlled interactions, the behaviour by both mice was observed and recorded in the above table. Each mouse was identified by a fur clip on either their right shoulder or right rump. A tick represents that that particular behaviour was observed. More ticks indicates more persistent behaviour. Aggressive behaviour was stopped by splitting the mice up to avoid injuries being sustained.

		Behaviour						
Interaction	Male	Sniffing	Approach	Retreat	Chasing	Escaping	Aggression	Avoidance
Day 1	A	✓✓	✓✓					
	B	✓✓	✓✓					
Day 2	A	✓✓	✓✓					
	B	✓✓	✓✓					
Day 3	A	✓✓	✓✓					
	B	✓✓	✓✓					

Supplementary material A: Recording of the behaviour observed in the third male/male pair.
During the controlled interactions, the behaviour by both mice was observed and recorded in the above table. Each mouse was identified by a fur clip on either their right shoulder or right rump. A tick represents that that particular behaviour was observed. More ticks indicates more persistent behaviour. Aggressive behaviour was stopped by splitting the mice up to avoid injuries being sustained.

Preparation of labelled diet

To make 1 kg labelled diet:

1. Identify the weight of labelled amino acid required to be added to the diet to achieve a 50 % labelled diet.
2. Rinse all glassware & blender in RO water.
3. In a large beaker, dissolve the appropriate amount of labelled amino acid in 1L RO water (equivalent volume for the weight of diet being prepared).
4. Add 1 kg unlabelled diet (Certified Rodent Diet 5002) and stir.
5. Leave for 2 hours then add a further 125 ml RO water and stir aggressively.
6. Leave for a further 1 hour, then add mix to blender, using a further 125 ml RO water to rinse beaker.
7. Blend for a minimum of 10 minutes until the mixture is a thick paste.
8. Alternate between high and low speeds, scraping sides occasionally.
9. Transfer food to a piping bag.
10. Attach baking paper (shiny side up) to the dehydrator trays.
11. Pipe food onto the trays and score pellet size lines across food.
12. Place in dehydrator for approximately 48 hours at 40°C.
13. If making two labelled diets at the same time, be sure to use two sets of clearly-labelled beakers & spatulas in order to prevent cross-contamination. Ensure that the blender is washed thoroughly with RO water in between blending each diet.

Supplementary material B: Full protocol for the preparation of the labelled diets used throughout the communal nursing study.

Supplementary material B: Assessing the rate of label incorporation

Spreadsheet where MUP peptide intensities were recorded, along with the intensities of the corresponding labelled peptides, for the mice fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

ID:	Sample no.	MUP peptide	Light peptide				Heavy lysine peptide				IH/(IH+IL)	Sample average RIA	Time of sampling (hours)	Day of sampling
			1	2	3	Sum	1	2	3	Sum				
10910	34081	2471/2477	122541	170550	120315	413406	0	0	0	0	0	0	0	0
		1823/1829	16631	17733	10348	44712	0	0	0	0	0			
10911	34082	2471/2477	909	1206	929	3044	0	0	0	0	0	0	0	0
		1823/1829	1896	1812	1096	4804	0	0	0	0	0			
10911	34085	2471/2477	62118	84985	64830	211933	6405	6318	4792	17515	0.076335	0.094474	6	1
		1823/1829	6602	5985	4016	16603	807	775	525	2107	0.112614			
10910	34086	2471/2477	14105	19791	15717	49613	2332	2374	1870	6576	0.117034	0.142572	6	1
		1823/1829	1700	1488	944	4132	289	295	251	835	0.16811			
10911	34089	2471/2477	23522	32792	29011	85325	10797	12308	10669	33774	0.283579	0.30141	12	1
		1823/1829	5556	5373	4175	15104	2788	2491	1804	7083	0.319241			
10910	34090	2471/2477	294815	370261	309143	974219	51366	52122	36842	140330	0.125907	0.114549	12	1
		1823/1829	80312	87401	53223	220936	9277	9776	6369	25422	0.103191			
10911	34093	2471/2477	12157	18739	16394	47290	8149	10188	9579	27916	0.371194	0.37292	18	1
		1823/1829	5218	5972	4033	15223	3262	3609	2249	9120	0.374646			
10910	34094	2471/2477	52074	72712	57327	182113	10782	11936	8552	31270	0.146544	0.152512	18	1
		1823/1829	8744	9696	6060	24500	1603	1819	1192	4614	0.15848			
10911	34097	2471/2477	12623	19597	17390	49610	7060	9233	8602	24895	0.334139	0.3576	30	2
		1823/1829	2491	2890	2175	7556	1686	1515	1451	4652	0.381062			
10910	34098	2471/2477	19986	31774	28644	80404	11587	15008	12729	39324	0.328444	0.338769	30	2
		1823/1829	4493	4797	3540	12830	2510	2483	1888	6881	0.349094			
10911	34101	2471/2477	50529	56632	44000	151161	28841	26895	20964	76700	0.336609	0.388774	36	2

		1823/1829	3132	4116	3278	10526	2988	3044	2270	8302	0.440939			
10910	34102	2471/2477	50313	76950	79221	206484	53409	67825	67466	188700	0.477499	0.471759	36	2
		1823/1829	7883	9378	7316	24577	6967	8019	6463	21449	0.466019			
10911	34105	2471/2477	17490	26699	27728	71917	17323	21796	19700	58819	0.449907	0.440048	42	2
		1823/1829	2103	2389	1850	6342	1546	1841	1401	4788	0.430189			
10910	34106	2471/2477	60735	87681	91461	239877	63494	80633	79638	223765	0.482625	0.481474	42	2
		1823/1829	13011	16586	12370	41967	13165	14456	11168	38789	0.480323			
10911	34107	2471/2477	30202	47970	46950	125122	22849	28195	25624	76668	0.37994	0.361442	54	3
		1823/1829	2944	3748	2742	9434	1727	1837	1360	4924	0.342945			
10910	34108	2471/2477	65061	96975	96640	258676	49400	60740	54517	164657	0.388954	0.378855	54	3
		1823/1829	11596	12788	9914	34298	6795	7915	5326	20036	0.368756			
10910	34113	2471/2477	23431	38871	42430	104732	30934	42386	42027	115347	0.524116	0.510803	66	3
		1823/1829	7669	9406	7251	24326	7688	9071	7324	24083	0.49749			
10911	34114	2471/2477	15328	25292	28331	68951	18457	25572	24095	68124	0.496983	0.484488	66	3
		1823/1829	3775	3810	3170	10755	3175	3339	3100	9614	0.471992			
10911	34117	2471/2477	39812	63603	63462	166877	31505	42551	38302	112358	0.402378	0.392857	78	4
		1823/1829	8267	10501	7632	26400	5482	6221	4708	16411	0.383336			
10910	34118	2471/2477	940	1658	1613	4211	762	812	732	2306	0.353844	0.381216	78	4
		1823/1829	260	376	273	909	186	222	220	628	0.408588			
10911	34129	2471/2477	39861	65646	66034	171541	43364	58742	55696	157802	0.479142	0.475258	90	4
		1823/1829	17328	21310	15690	54328	16651	18381	13412	48444	0.471374			
10910	34130	2471/2477	13052	20232	21750	55034	15644	21008	21071	57723	0.511924	0.509968	90	4
		1823/1829	5907	6361	5630	17898	6355	6580	5546	18481	0.508013			

10911	34135	2471/2477	34993	60957	57969	153919	24974	33302	30639	88915	0.366155	0.373271	102	5
		1823/1829	5577	6598	5352	17527	3592	3876	3292	10760	0.380387			
10910	34136	2471/2477	118819	183997	170223	473039	80445	103414	89349	273208	0.366109	0.353545	102	5
		1823/1829	12513	14951	10226	37690	7039	7396	5066	19501	0.34098			
10911	34137	2471/2477	44254	70280	73637	188171	51389	68247	69103	188739	0.500753	0.502237	114	5
		1823/1829	7082	9134	6860	23076	7892	8998	6532	23422	0.503721			
10910	34138	2471/2477	6131	10091	10359	26581	6713	8727	8220	23660	0.47093	0.47471	114	5
		1823/1829	2354	2706	2153	7213	2322	2360	1936	6618	0.47849			

Supplementary material B: Assessing the rate of label incorporation

Spreadsheet where MUP peptide intensities were recorded, along with the intensities of the corresponding labelled peptides, for the mice fed [$^{2}H_{8}$] valine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

ID:	Sample no.	MUP peptide	Light valine peptide				Heavy valine peptide				IH/(IH+IL)	Sample average RIA	Time of sampling (hours)	Day of sampling
			1	2	3	Sum	1	2	3	Sum				
10903	34079	2009/2017	2698	3090	1875	7663	0	0	0	0	0	0	0	0
		1567/1575	1682	1272	499	3453	0	0	0	0	0			
10904	34080	2009/2017	1677	1339	1242	4258	0	0	0	0	0	0	0	0
		1567/1575	234	192	132	558	0	0	0	0	0			
10904	34083	2009/2017	13954	17276	10361	41591	1786	1602	1049	4437	0.096398	0.133255	6	1
		1567/1575	1571	1622	661	3854	337	239	214	790	0.170112			
10903	34084	2009/2017	32586	42986	27820	103392	9666	8657	4831	23154	0.182969	0.228568	6	1
		1567/1575	2736	2500	1086	6322	991	767	630	2388	0.274168			
10904	34087	2009/2017	9495	11042	6550	27087	2097	1804	955	4856	0.152021	0.191343	12	1
		1567/1575	1399	965	501	2865	211	363	285	859	0.230666			
10903	34088	2009/2017	7248	9105	5812	22165	5516	7446	5483	18445	0.454198	0.443219	12	1
		1567/1575	1072	783	445	2300	472	771	508	1751	0.432239			
10904	34091	2009/2017	4343	5087	3285	12715	1993	1329	669	3991	0.238896	0.213103	18	1
		1567/1575	1270	1137	526	2933	328	199	149	676	0.18731			
10903	34092	2009/2017	4194	4865	2983	12042	4971	3671	1663	10305	0.461136	0.410554	18	1
		1567/1575	1317	904	526	2747	633	612	300	1545	0.359972			
10904	34095	2009/2017	11301	12700	7683	31684	1698	1698	953	4349	0.120695	0.193788	30	2
		1823/1831	4163	4134	2842	11139	1584	1730	741	4055	0.266882			
10903	34096	2009/2017	1673	2167	1246	5086	1004	797	403	2204	0.302332	0.339845	30	2
		1567/1575	515	438	202	1155	289	274	137	700	0.377358			
10903	34099	2009/2017	12596	17549	11735	41880	12790	12065	9377	34232	0.449758	0.431776	36	2
		1567/1575	1806	1351	821	3978	1345	973	490	2808	0.413793			
10904	34100	2009/2017	2340	3272	1884	7496	1585	1254	939	3778	0.335107	0.335107	36	2
10904	34103	2009/2017	4780	5186	3330	13296	3196	3208	1513	7917	0.373215	0.373215	42	2
10903	34104	2009/2017	26039	35497	23888	85424	15932	13939	7254	37125	0.30294	0.401412	42	2
		1567/1575	1804	1628	900	4332	1001	1917	1412	4330	0.499885			

10903	34109	2009/2017	10282	14476	8959	33717	6498	5904	2975	15377	0.313215	0.349735	54	3
		1567/1575	922	852	414	2188	685	485	207	1377	0.386255			
10904	34110	1596/1604	1235	1060	629	2924	599	366	366	1331	0.312808	0.312808	54	3
10904	34112	2009/2017	18090	22075	13979	54144	18013	16435	9290	43738	0.446844	0.448321	66	3
		1567/1575	4053	3755	1897	9705	3840	2577	1517	7934	0.449799			
10904	34115	2009/2017	16835	1902	15835	34572	3949	12589	7326	23864	0.408378	0.409057	78	4
		1567/1575	3545	3220	2439	9204	2457	2215	1717	6389	0.409735			
10903	34116	2009/2017	9965	12776	8137	30878	8575	7220	3577	19372	0.385512	0.402363	78	4
		1567/1575	1168	1230	661	3059	975	686	547	2208	0.419214			
10904	34127	2009/2017	11326	13916	9146	34388	11737	9097	5186	26020	0.430738	0.418962	90	4
		1567/1575	10921	9452	4457	24830	8659	5841	2555	17055	0.407186			
10903	34128	2009/2017	24222	34298	23942	82462	24416	19832	10063	54311	0.397089	0.427589	90	4
		1567/1575	2590	2281	1439	6310	2474	1838	1022	5334	0.45809			
10904	34133	2009/2017	12991	16927	10444	40362	7055	5769	3027	15851	0.281981	0.374293	102	5
		1567/1575	1533	403	931	2867	1042	843	623	2508	0.466605			
10903	34134	2009/2017	2074	2918	2036	7028	1360	1172	615	3147	0.309287	0.370733	102	5
		1567/1575	676	653	425	1754	504	561	270	1335	0.432179			
10903	34139	2009/2017	1799	2235	1660	5694	2170	1743	1084	4997	0.467402	0.467402	114	5
10904	34140	2009/2017	17195	20287	13034	50516	28040	20682	10204	58926	0.538422	0.47655	114	5
		1567/1575	3312	2924	1588	7824	2724	1870	949	5543	0.414678			

Supplementary B: Proof-of-principle experiment

Spreadsheet where MUP peptide intensities were recorded, along with the intensities of the corresponding labelled peptides, for the mice fed [$^{12}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

		Lysine peptide	Light lysine peptide intensity				Heavy lysine peptide intensity				Valine peptide	Light valine peptide intensity				Heavy valine peptide intensity					
Sample number	Day of labelling	Light/heavy MUP peptide (Da)	1	2	3	Sum	1	2	3	Sum	Light/heavy MUP peptide (Da)	1	2	3	Sum	1	2	3	Sum	Lysine IH/(IH+IL)	Valine IH/(IH+IL)
10978	0	2471/2477	20980	28152	21552	70684	0	0	0	0	2009/2017	2119	2712	2335	7166	0	0	0	0	0	0
		2009/2015	2119	2712	2335	7166	0	0	0	0	1567/1575	576	307	284	1167	0	0	0	0	0	0
	1	2471/2477	10626	16642	16669	43937	8050	9442	8333	25825	2009/2017	2087	2142	1711	5940	0	0	0	0	0.370187	0
		2009/2015	2087	2142	1711	5940	1273	1470	1215	3958	1567/1575	369	442	349	1160	0	0	0	0	0.399879	0
	2	2471/2477	6466	10042	11728	28236	6831	8487	8752	24070	2009/2017	806	1029	905	2740	0	0	0	0	0.460177	0
		2009/2015	806	1029	905	2740	751	743	608	2102	1567/1575	0	0	0	0	0	0	0	0	0.434118	0
	3	2471/2477	11190	18382	18353	47925	10845	13520	13170	37535	2009/2017	1363	1796	1418	4577	0	0	0	0	0.439211	0
		2009/2015	1363	1796	1418	4577	1334	1273	1022	3629	1567/1575	233	217	207	657	0	0	0	0	0.442237	0
	4	2471/2477	15188	24314	23859	63361	11779	14653	13638	40070	2009/2017	3364	3637	2699	9700	0	0	0	0	0.387408	0
		2009/2015	3364	3637	2699	9700	2120	2383	2212	6715	1567/1575	586	671	552	1809	0	0	0	0	0.409077	0
	5	2471/2477	42706	65160	66975	174841	37714	45357	41738	124809	2009/2017	6245	6998	5324	18567	0	0	0	0	0.416516	0
		2009/2015	6245	6998	5324	18567	4470	5789	4787	15046	1567/1575	1392	1289	1136	3817	0	0	0	0	0.447624	0
10987	0	2471/2477	187883	227002	188254	603139	0	0	0	0	2009/2017	40871	45691	28155	114717	0	0	0	0	0	0
		2009/2015	40871	45691	28155	114717	0	0	0	0	1567/1575	7284	5448	2941	15673	0	0	0	0	0	0
	2	2471/2477	36959	56954	55373	149286	30715	38115	35601	104431	2009/2017	10268	12251	8733	31252	0	0	0	0	0.411604	0
		2009/2015	10268	12251	8733	31252	7361	7957	5871	21189	1567/1575	908	917	701	2526	0	0	0	0	0.404054	0
	3	2471/2477	70254	108187	108396	286837	62964	81331	77991	222286	2009/2017	33383	41509	31335	106227	0	0	0	0	0.436606	0
		2009/2015	33383	41509	31335	106227	25872	31335	22433	79640	1567/1575	3906	3882	3216	11004	0	0	0	0	0.428478	0
	4	2471/2477	14780	21617	21034	57431	10242	13628	12089	35959	2009/2017	11546	13956	9582	35084	0	0	0	0	0.385041	0

		2009/2015	11546	13956	9582	35084	7290	8495	5913	21698	1567/1575	3019	2945	2223	8187	0	0	0	0	0.382128	0
	5	2471/2477	51385	79893	77698	208976	42680	55351	52378	150409	2009/2017	3064 8	3806 2	2863 1	97341	0	0	0	0	0.418518	0
		2009/2015	30648	38062	28631	97341	22358	27034	20242	69634	1567/1575	2843	2454	1908	7205	0	0	0	0	0.417032	0
10865	0	2471/2477	10303 2	12828 6	10164 5	33296 3	0	0	0	0	2009/2017	2926 3	3308 1	1957 4	81918	0	0	0	0	0	0
		2009/2015	29263	33081	19574	81918	0	0	0	0	1567/1575	4147	3147	1276	8570	0	0	0	0	0	0
	1	2471/2477	77845	11452 7	11287 2	30524 4	65001	80880	76170	222051	2009/2017	3100 1	3936 8	2822 3	98592	0	0	0	0	0.421113	0
		2009/2015	31001	39368	28223	98592	29504	33776	26231	89511	1567/1575	5786	5652	4121	15559	0	0	0	0	0.475862	0
	2	2471/2477	28558	47965	49984	126507	37007	47526	46830	131363	2009/2017	4685 5	5751 4	4420 4	148573	0	0	0	0	0.509416	0
		2009/2015	46855	57514	44204	148573	48483	54372	41825	144680	1567/1575	4566	4025	3549	12140	0	0	0	0	0.493362	0
	3	2471/2477	44160	65994	65857	176011	41696	51513	47576	140785	2009/2017	2115 3	2468 1	1764 3	63477	0	0	0	0	0.444403	0
		2009/2015	21153	24681	17643	63477	16555	18487	13680	48722	1567/1575	2297	2408	1832	6537	0	0	0	0	0.434246	0
	4	2471/2477	160814	238985	220618	620417	112527	142840	144721	400088	2009/2017	3362 1	4232 7	3191 1	107859	0	0	0	0	0.392049	0
		2009/2015	33621	42327	31911	107859	32169	39629	35869	107667	1567/1575	4249	4011	2818	11078	0	0	0	0	0.499555	0
	5	2471/2477	25550	40417	42084	108051	32085	42065	42224	116374	2009/2017	1386 9	1781 1	1474 3	46423	0	0	0	0	0.518543	0
		2009/2015	13869	17811	14743	46423	15650	18118	14278	48046	1567/1575	969	960	940	2869	0	0	0	0	0.50859	0
10827	0	2471/2477	56907	75626	55238	187771	0	0	0	0	2009/2017	6380	7121	4606	18107	0	0	0	0	0	0
		2009/2015	6380	7121	4606	18107	0	0	0	0	1567/1575	1889	1606	858	4353	0	0	0	0	0	0
	1	2471/2477	94966	138665	133906	367537	79652	99629	92546	271827	2009/2017	1783 5	2232 8	1619 5	56358	0	0	0	0	0.425152	0
		2009/2015	17835	22328	16195	56358	13612	16299	11395	41306	1567/1575	1612	1390	1240	4242	0	0	0	0	0.42294	0
	2	2471/2477	29263	45206	45596	120065	27495	34788	33316	95599	2009/2017	5273	6322	4384	15979	0	0	0	0	0.443278	0
		2009/2015	5273	6322	4384	15979	4123	4709	3353	12185	1567/1575	727	735	597	2059	0	0	0	0	0.432645	0
	3	2471/2477	42569	63331	66698	172598	46788	59222	57937	163947	2009/2017	1789 3	2178 5	1659 1	56269	0	0	0	0	0.487147	0

		2009/2015	17893	21785	16591	56269	17559	21085	14973	53617	1567/1575	2475	2214	2002	6691	0	0	0	0	0.487933	0
	4	2471/2477	64185	94788	88620	247593	46537	58420	50494	155451	2009/2017	19549	24646	16299	60494	0	0	0	0	0.385692	0
		2009/2015	19549	24646	16299	60494	12338	14166	9218	35722	1567/1575	2231	2068	1312	5611	0	0	0	0	0.371269	0
	5	2471/2477	48253	76273	82839	207365	64349	83414	85617	233380	2009/2017	21588	26208	20478	68274	0	0	0	0	0.529513	0
		2009/2015	21588	26208	20478	68274	23969	27369	20456	71794	1567/1575	2418	2322	2161	6901	0	0	0	0	0.512565	0
10954	0	2471/2477	208610	279038	236864	724512	0	0	0	0	2009/2017	35203	42452	26401	104056	0	0	0	0	0	0
		2009/2015	35203	42452	26401	104056	0	0	0	0	1567/1575	13131	10849	7154	31134	0	0	0	0	0	0
	1	2471/2477	99455	148139	145498	393092	89825	112955	106746	309526	2009/2017	40666	49304	36339	126309	0	0	0	0	0.440532	0
		2009/2015	40666	49304	36339	126309	32729	38082	27430	98241	1567/1575	1628	1779	1511	4918	0	0	0	0	0.437502	0
	2	2471/2477	62854	96494	96104	255452	51855	65908	61553	179316	2009/2017	20183	23960	17038	61181	0	0	0	0	0.412441	0
		2009/2015	20183	23960	17038	61181	13788	15906	11174	40868	1567/1575	1959	1803	1378	5140	0	0	0	0	0.400474	0
	3	2471/2477	16022	24954	25056	66032	15566	20125	18422	54113	2009/2017	7800	10361	7471	25632	0	0	0	0	0.450397	0
		2009/2015	7800	10361	7471	25632	6969	7934	5821	20724	1567/1575	2205	2240	1896	6341	0	0	0	0	0.447062	0
	4	2471/2477	68058	103674	106268	278000	71169	89285	87195	247649	2009/2017	26969	33899	26431	87299	0	0	0	0	0.47113	0
		2009/2015	26969	33899	26431	87299	23937	27815	20705	72457	1567/1575	2949	2883	2281	8113	0	0	0	0	0.453548	0
	5	2471/2477	37831	56367	51356	145554	23921	30742	26167	80830	2009/2017	16805	20220	14380	51405	0	0	0	0	0.357048	0
		2009/2015	16805	20220	14380	51405	9340	10576	7614	27530	1567/1575	2357	2112	1365	5834	0	0	0	0	0.348768	0

Supplementary B: Proof-of-principle experiment

Spreadsheet where MUP peptide intensities were recorded, along with the intensities of the corresponding labelled peptides, for the mice fed [$^{2}H_{8}$] valine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

		Lysine/valine peptide	Light valine/lysine peptide intensity				Heavy lysine peptide intensity				Heavy valine peptide intensity					
Sample number	Day of labelling	Light/heavy MUP peptide (Da)	1	2	3	Sum	1	2	3	Sum	1	2	3	Sum	Lysine IH/(IH+IL)	Valine IH/(IH+IL)
10977	0	2009/2017	37361	44346	27096	108803	0	0	0	0	0	0	0	0	0	0
		1567/1575	2461	2196	1858	6515	0	0	0	0	0	0	0	0	0	0
	1	2009/2017	5199	6417	4456	16072	0	0	0	0	3736	2977	1929	8642	0	0.34968
		1567/1575	999	861	623	2483	0	0	0	0	662	889	735	2286	0	0.479346
	2	2009/2017	2646	3193	2070	7909	0	0	0	0	1302	2321	2142	5765	0	0.421603
	3	2009/2017	17868	21295	13705	52868	0	0	0	0	17873	15304	8781	41958	0	0.442474
		1567/1575	1555	1568	1338	4461	0	0	0	0	1173	1480	1323	3976	0	0.471258
	4	2009/2017	978	2070	2183	5231	0	0	0	0	710	1257	1119	3086	0	0.371047
		1567/1575	184	201	158	543	0	0	0	0	176	322	153	651	0	0.545226
	5	2009/2017	11820	14525	9944	36289	0	0	0	0	13018	11390	6180	30588	0	0.457377
		1567/1575	1644	1425	1029	4098	0	0	0	0	1906	1189	944	4039	0	0.496375
10986	0	2009/2017	422	538	404	1364	0	0	0	0	0	0	0	0	0	0
		1567/1575	399	402	401	1202	0	0	0	0	0	0	0	0	0	0
	1	2009/2017	5346	5801	4281	15428	0	0	0	0	5184	4168	3025	12377	0	0.445136
		1567/1575	2416	2452	2179	7047	0	0	0	0	2674	2283	1855	6812	0	0.491522
	2	2009/2017	4523	5322	3430	13275	0	0	0	0	3545	2973	1743	8261	0	0.38359
		1567/1575	715	725	538	1978	0	0	0	0	654	536	498	1688	0	0.460447
	3	2009/2017	12908	14741	9124	36773	0	0	0	0	10031	8393	4737	23161	0	0.386442
		1567/1575	3774	3347	2093	9214	0	0	0	0	2752	1986	1609	6347	0	0.407879
	5	2009/2017	15255	18922	11824	46001	0	0	0	0	10234	9225	5255	24714	0	0.349487
		1567/1575	3747	3569	2017	9333	0	0	0	0	3120	2177	1517	6814	0	0.421998
10840	0	2009/2017	19361	21942	13248	54551	0	0	0	0	0	0	0	0	0	0
		1567/1575	4615	3574	2009	10198	0	0	0	0	0	0	0	0	0	0
	1	2009/2017	4395	4922	3302	12619	0	0	0	0	4310	3648	2268	10226	0	0.447625
		1567/1575	1854	1870	1247	4971	0	0	0	0	1709	1508	1095	4312	0	0.464505

	2	2009/2017	15412	18620	11885	45917	0	0	0	0	18699	15969	9451	44119	0	0.490015
		1567/1575	2539	1833	1033	5405	0	0	0	0	2527	1540	993	5060	0	0.483516
	3	2009/2017	20630	26597	16172	63399	0	0	0	0	20712	18432	10310	49454	0	0.438216
		1567/1575	2870	2380	1415	6665	0	0	0	0	2477	1781	1290	5548	0	0.45427
	4	2009/2017	10513	12766	7932	31211	0	0	0	0	11180	9646	5628	26454	0	0.458753
		1567/1575	2148	1946	872	4966	0	0	0	0	1983	1546	776	4305	0	0.464351
	5	2009/2017	26678	32729	20737	80144	0	0	0	0	28390	25811	14740	68941	0	0.462427
		1567/1575	4133	3590	1941	9664	0	0	0	0	4376	2929	1768	9073	0	0.484229
10829	0	2009/2017	8096	8923	5391	22410	0	0	0	0	0	0	0	0	0	0
		1567/1575	1724	1254	746	3724	0	0	0	0	0	0	0	0	0	0
	1	2009/2017	838	776	620	2234	0	0	0	0	607	620	376	1603	0	0.417774
		1567/1575	358	348	227	933	0	0	0	0	330	277	192	799	0	0.461316
	2	2009/2017	10464	13026	7605	31095	0	0	0	0	8601	7865	4141	20607	0	0.398573
		1567/1575	1413	1346	664	3423	0	0	0	0	1124	747	503	2374	0	0.409522
	3	2009/2017	6787	8741	5695	21223	0	0	0	0	3686	3406	1650	8742	0	0.29174
		1567/1575	1050	1046	618	2714	0	0	0	0	1203	881	484	2568	0	0.486179
	4	2009/2017	9187	11674	7674	28535	0	0	0	0	7380	7006	3756	18142	0	0.388671
		1567/1575	1236	958	538	2732	0	0	0	0	956	750	453	2159	0	0.441423
	5	2009/2017	10745	12655	8424	31824	0	0	0	0	9002	8543	4350	21895	0	0.407584
		1567/1575	2086	1754	966	4806	0	0	0	0	1646	1088	815	3549	0	0.424776
10952	0	2009/2017	29903	33881	20774	84558	0	0	0	0	0	0	0	0	0	0
		1567/1575	2487	2043	1397	5927	0	0	0	0	0	0	0	0	0	0
	1	2009/2017	22505	26292	16456	65253	0	0	0	0	22707	19417	11049	53173	0	0.448998
		1567/1575	2120	1647	911	4678	0	0	0	0	1760	1181	891	3832	0	0.450294
	2	2009/2017	20860	25291	15974	62125	0	0	0	0	20020	17257	10353	47630	0	0.433967
		1567/1575	3083	2390	1287	6760	0	0	0	0	2729	1959	1286	5974	0	0.469138
	3	2009/2017	6201	6986	5004	18191	0	0	0	0	3571	7170	6730	17471	0	0.489905
		1567/1575	1247	939	694	2880	0	0	0	0	1362	1168	681	3211	0	0.527171
	4	2009/2017	15778	19513	13050	48341	0	0	0	0	19188	16849	10075	46112	0	0.4882

		1567/1575	2776	2322	1573	6671	0	0	0	0	2953	2114	1302	6369	0	0.48842
	5	2009/2017	35203	42452	26641	104296	0	0	0	0	31222	27108	14512	72842	0	0.411216
		1567/1575	3914	3426	1901	9241	0	0	0	0	3513	2719	1536	7768	0	0.456699

Sample no.	Day of sample	MUP peptide	Light peptide				Heavy lysine peptide				IH/(IH+IL)	Day average
			1	2	3	Sum	1	2	3	Sum		
34965	0	2471/2477	266641	366065	304321	937027	80668	93988	77120	251776	0.21179	0.283576
34966	0	2471/2477	74121	107889	90682	272692	30131	34365	28102	92598	0.253492	
34967	0	2471/2477	186806	266038	248873	701717	131180	160735	148202	440117	0.385447	
34974	1	2471/2477	37276	59851	67516	164643	49733	61958	64321	176012	0.516687	0.470132
34975	1	2471/2477	2055	2986	2443	7484	2275	2011	2056	6342	0.458701	
34976	1	2471/2477	177839	271406	273638	722883	166053	200401	190117	556571	0.435007	
35283	2	2471/2477	126450	202947	227332	556729	165922	209927	216300	592149	0.515415	0.433976
35284	2	2471/2477	4312	6068	5700	16080	3532	4611	4607	12750	0.442248	
35285	2	2471/2477	124387	183206	174204	481797	77936	93940	81072	252948	0.344266	
35292	3	2471/2477	200646	315405	327221	843272	220807	274256	271354	766417	0.476127	0.495727
35293	3	2471/2477	33857	55044	61009	149910	44591	56652	59089	160332	0.516797	
35294	3	2471/2477	9094	16446	17469	43009	11508	15466	15058	42032	0.494256	
35301	4	2471/2477	34688	57733	65308	157729	51617	64813	65807	182237	0.536045	0.48667
35302	4	2471/2477	3831	4478	4452	12761	3211	3457	3249	9917	0.437296	
35309	5	2471/2477	39326	61200	70654	171180	56303	72035	76036	204374	0.544193	0.535173
35310	5	2471/2477	28430	48545	54774	131749	40136	51811	54345	146292	0.526153	
35317	6	2471/2477	78347	127933	144965	351245	109003	140628	146386	396017	0.529957	0.529702
35318	6	2471/2477	12202	20870	24361	57433	17911	22536	24174	64621	0.529446	

Supplementary material B: Milk labelling pilot

Spreadsheet where MUP peptide intensities were recorded, along with the intensities of the corresponding labelled peptides, for the mice fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

ID:	Day of sample	Milk peptide	Light peptide				Heavy lysine peptide				IH/(IH+IL)	Day average
			1	2	3	Sum	1	2	3	Sum		
34963	0	2433	5653	7101	6339	19093	0	0	0	0	0	0
34968	1	2433	6504	8916	8309	23729	6534	6971	6596	20101	0.458613	0.409958
34969	1	2433	1891	2648	2700	7239	1303	1335	1457	4095	0.361302	
34977	2	2433	302	414	384	1100	285	344	354	983	0.471916	0.428103
34978	2	2433	31381	42979	39651	114011	22062	26675	22422	71159	0.38429	
35286	3	2433	1474	2153	1897	5524	1204	1407	1299	3910	0.414458	0.413974
35287	3	2433	1068	1309	1023	3400	845	877	675	2397	0.41349	
35295	4	2433	6418	8812	8274	23504	6487	6934	6572	19993	0.459641	0.48362
35296	4	2433	392	417	422	1231	461	341	467	1269	0.5076	
35303	5	2433	4469	5813	5414	15696	3800	4844	3832	12476	0.442851	0.466156
35304	5	2433	2805	3712	3584	10101	2796	3603	3285	9684	0.489462	
35311	6	2433	1332	2086	1903	5321	1503	1972	1966	5441	0.505575	0.491051
35312	6	2433	22501	32617	30774	85892	24154	29256	24779	78189	0.476527	

Supplementary material B: Milk labelling pilot

Spreadsheet where the lactation elevation protein peptide intensities in pup stomach contents were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

ID:	Day of sample	Urine peptide	Light peptide				Heavy lysine peptide				IH/(IH+IL)	Day average
			1	2	3	Sum	1	2	3	Sum		
34972	1	2000	11971	12956	7057	31984	669	724	586	1979	0.058269	0.100266
34973	1	2000	7649	8968	5836	22453	964	1297	1463	3724	0.142262	
35281	2	2000	8558	11517	7610	27685	999	1202	895	3096	0.100582	0.083762
35282	2	2000	11388	13167	7851	32406	864	796	665	2325	0.066943	
35290	3	2000	4936	5675	3690	14301	867	1079	753	2699	0.158765	0.180883
35291	3	2000	704	758	662	2124	141	201	199	541	0.203002	
35300	4	2000	24365	29195	18700	72260	5774	6308	4222	16304	0.184093	0.184093
35308	5	2000	5407	6718	4126	16251	1428	1392	1307	4127	0.202522	0.202522
35316	6	2000	1549	2131	2071	5751	763	944	517	2224	0.278871	0.278871

Supplementary material B: Milk labelling pilot

Spreadsheet where the serum albumin peptide intensities in pup urine were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

Supplementary material B: Milk labelling pilot

Spreadsheet where the peptide intensities of the top 20 pup liver proteins were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

		Light peptide intensity				Heavy peptide intensity				
Sample ID	Protein/peptide	1	2	3	Sum	1	2	3	Sum	IH/(IH+IL)
	P02089									
35313	912.47	6896	3550	1090	11536	1752	816	256	2824	0.196657
	1091.6	1044	604	455	2103	215	135	85	435	0.171395
	1294.6	95194	70309	30088	195591	18784	15406	6461	40651	0.172074
35314	912.47	1251464	523167	168050	1942681	204000	102350	28506	334856	0.147025
	1091.6	497551	288602	102659	888812	81747	44500	15293	141540	0.137371
	1294.6	739675	522655	211779	1474109	127890	89967	41293	259150	0.149516
	P12710									
35313	1788.8	32935	32864	24678	90477	11650	12588	7070	31308	0.257076
	1196.6	154994	112388	48951	316333	56410	34937	16712	108059	0.254621
	892.5	133783	64928	23603	222314	44010	20682	6680	71372	0.243021
35314	1788.8	21185	22746	14189	58120	7745	8398	6287	22430	0.278461
	1196.6	117150	84373	37516	239039	43542	27282	12654	83478	0.258833
	892.5	126536	61288	18962	206786	43282	17808	5606	66696	0.243877
	B1AXW4									
35313	1263.7	32557	25415	10942	68914	14412	9331	4731	28474	0.292377
	1164.5	36510	24371	12290	73171	15097	9842	4310	29249	0.285579
	954.5	42708	24216	7630	74554	18321	9070	3279	30670	0.291473
35314	1263.7	23454	20709	8816	52979	9530	6356	2730	18616	0.260018
	1164.5	25526	19348	9351	54225	10696	6638	3182	20516	0.274495
	954.5	31793	17306	6369	55468	12294	5942	2346	20582	0.270638
	P26443									
35313	1583.8	16780	15915	9468	42163	5677	3718	2449	11844	0.219305
	716.4	18548	6467	1439	26454	6102	1779	617	8498	0.243133
	1125.5	40049	24615	9128	73792	11846	6733	2578	21157	0.222825
35314	1583.8	8157	6430	3114	17701	2419	2125	914	5458	0.235675

	716.4	12770	5287	1528	19585	4424	1518	389	6331	0.244289
	1125.5	27618	16723	6099	50440	8777	4678	2217	15672	0.237052
	Q9QXD6									
35313	1328.7	26474	19890	9719	56083	10458	7257	3447	21162	0.273959
	1608.8	5730	6571	2943	15244	2542	2194	1242	5978	0.281689
	1292.7	8981	7486	2920	19387	3821	2695	1523	8039	0.293116
35314	1328.7	24393	20238	9035	53666	9405	6994	3687	20086	0.272345
	1608.8	56579	53041	29871	139491	24330	19711	10508	54549	0.281122
	1292.7	9940	6626	3420	19986	2854	2809	1119	6782	0.253362
	P09103									
35313	1409.7	4354	4107	3107	11568	3997	2150	1196	7343	0.388293
	1216.6	6271	4362	2403	13036	2469	1830	626	4925	0.274205
	1052.6	8813	5767	1974	16554	3451	1940	771	6162	0.271263
35314	1409.7	5234	4286	2369	11889	2975	2285	1481	6741	0.361836
	1216.6	2460	1665	860	4985	724	785	415	1924	0.278477
	1052.6	10168	5579	2207	17954	3592	1786	811	6189	0.256348
	P99027									
35313	1242.6	7983	6151	4543	18677	2384	2073	844	5301	0.221078
	1772.9	7605	9085	4416	21106	2579	2057	385	5021	0.192177
	2774.4	5414	5104	3814	14332	859	1795	1580	4234	0.228051
35314	1242.6	6855	5587	3065	15507	2260	2453	818	5531	0.262905
	1772.9	10054	10417	6387	26858	2909	3279	2186	8374	0.237682
	2774.4	4181	5938	6500	16619	2213	1747	1126	5086	0.234324
	Q8C196									
35313	1723.8	12075	11037	7029	30141	4200	3535	2297	10032	0.24972
	1217.6	24736	19487	8922	53145	8399	6229	2878	17506	0.247781
	1163.7	15216	11905	4817	31938	5032	4140	1099	10271	0.243337
35314	1723.8	10302	10980	5857	27139	4359	3837	2950	11146	0.291132
	1217.6	22841	17797	9175	49813	9711	6067	3768	19546	0.281809

	1163.7	15655	12945	4790	33390	6021	4507	1601	12129	0.26646
	P54869									
35313	1053.5	60108	33826	15322	109256	20202	11206	5071	36479	0.25031
	878.4	66232	35824	12878	114934	23419	10731	4324	38474	0.250795
	997.5	100355	57740	20816	178911	34524	17479	6998	59001	0.247995
35314	1053.5	34321	23874	11082	69277	11635	7041	4042	22718	0.246948
	878.4	34602	16813	5438	56853	10455	5116	2124	17695	0.237364
	997.5	42460	24242	8368	75070	15673	7599	3976	27248	0.266307
	P32020									
35313	1138.7	32644	20506	9123	62273	13615	9212	4052	26879	0.301496
	1046.5	22960	14729	6053	43742	12472	6898	6578	25948	0.372335
	1335.7	11716	8288	4022	24026	8916	6109	2870	17895	0.426874
35314	1138.7	15146	10747	4801	30694	7637	3894	2102	13633	0.307555
	1046.5	11058	6390	2574	20022	4108	3152	3175	10435	0.342614
	1335.7	8065	5788	3426	17279	6761	4390	2569	13720	0.442595
	P02088									
35313	912.8	214084	91347	31661	337092	38969	18729	5984	63682	0.158898
	1756.9	39081	39860	20455	99396	9143	8410	3706	21259	0.176197
	938.5	41463	20821	7294	69578	7561	3743	1390	12694	0.154293
35314	912.8	1251464	523167	168050	1942681	204000	102350	28506	334856	0.147025
	1756.9	517829	514769	249071	1281669	103773	93183	49095	246051	0.161058
	938.5	267512	134266	42758	444536	52423	25997	7016	85436	0.161209
	Q05816									
35313	1139.5	23711	13829	6698	44238	9334	5612	2012	16958	0.27711
	1055.7	19472	12290	4733	36495	8052	4095	1959	14106	0.278769
	2448.1	8355	11712	9541	29608	3752	5633	4153	13538	0.313772
35314	1139.5	15249	9050	3555	27854	4422	3050	1415	8887	0.241882
	1055.7	14911	8653	3086	26650	4340	2150	1332	7822	0.226909
	2448.1	5440	7482	6299	19221	2253	2853	2388	7494	0.280517

	P06151									
35313	1118.6	47965	28814	11426	88205	16709	12469	3748	32926	0.271821
	1250.7	38502	28759	13723	80984	14942	9614	4496	29052	0.264023
	1385.8	35689	28676	14422	78787	11408	10311	3948	25667	0.245725
35314	1118.6	39719	26408	9612	75739	12812	8469	2981	24262	0.242618
	1250.7	62172	72529	40956	175657	16904	12614	6404	35922	0.169781
	1385.8	24441	19131	10047	53619	8580	6914	3667	19161	0.263273
	B1ATY1									
35313	1161.6	29651	20855	9917	60423	12837	7093	3429	23359	0.278807
	923.6	12904	7675	2640	23219	4792	2532	1166	8490	0.267747
	1966.9	5584	7500	4506	17590	1973	2256	1382	5611	0.241843
35314	1161.6	25183	15573	8291	49047	9520	6782	2608	18910	0.278264
	923.6	12624	6757	2661	22042	4098	2527	878	7503	0.253952
	1966.9	5315	6600	4255	16170	2335	2412	1271	6018	0.271228
	Q64433									
35313	1285.7	14217	9016	4520	27753	4448	2464	1499	8411	0.232579
	1076.6	19934	12609	4171	36714	5390	3302	1767	10459	0.221716
	1044.6	19134	11525	3787	34446	4654	2657	1365	8676	0.201197
35314	1285.7	8942	6536	3418	18896	3849	2408	997	7254	0.2774
	1076.6	12185	7454	2694	22333	3804	2397	1256	7457	0.250319
	1044.6	12225	6704	3008	21937	3683	2325	984	6992	0.241695
	P16460									
35313	1186.7	5723	5008	2454	13185	2675	1646	1116	5437	0.291966
	936.5	27399	16304	6770	50473	13058	5633	2653	21344	0.2972
	806.5	24727	11439	3594	39760	9763	4834	1292	15889	0.285522
35314	1186.7	11701	7416	3557	22674	5647	3519	1287	10453	0.315543
	936.5	56400	30718	11443	98561	27478	12866	4438	44782	0.312411
	806.5	41075	17105	5712	63892	18715	9211	2607	30533	0.323357
	O35490									

35313	1721.9	70929	71946	42785	185660	28938	31999	20196	81133	0.304105
	975.5	40506	23644	9698	73848	17898	9820	4810	32528	0.305783
	2286.2	2568	2646	2702	7916	1245	963	897	3105	0.281735
35314	1721.9	70320	69821	43871	184012	29410	28846	15871	74127	0.287159
	975.5	29098	17616	8118	54832	12980	6726	3551	23257	0.297827
	2286.2	7963	7142	5581	20686	3116	3445	2306	8867	0.300037
	P07724									
35313	1149.6	45232	29825	12504	87561	21763	13373	5725	40861	0.318178
	972.6	48702	26569	8562	83833	23971	11548	4605	40124	0.323693
	761.4	17582	7858	2758	28198	9275	3945	1246	14466	0.339068
35314	1149.6	87335	57565	22792	167692	38137	22312	11513	71962	0.300275
	972.6	92934	52236	18212	163382	41141	24062	7153	72356	0.306934
	761.4	33197	15247	4664	53108	16805	5644	1653	24102	0.312162
	Q6PHC1									
35313	1960.9	4621	5790	3749	14160	1423	1785	1194	4402	0.237151
	944.5	12530	7311	3038	22879	5017	2954	1176	9147	0.285612
	2194.1	2825	4682	3989	11496	1416	1722	1329	4467	0.279835
35314	1960.9	4436	5624	3585	13645	2057	2138	1113	5308	0.280061
	944.5	12911	7608	3243	23762	3335	2267	1174	6776	0.221887
	2194.1	7875	10264	6981	25120	3015	4057	1833	8905	0.261719
	Q8QZT1									
35313	1024.5	22454	11485	4498	38437	8131	4887	1873	14891	0.279234
	829.5	25869	12253	4322	42444	8311	3404	1683	13398	0.239927
	1171.7	38431	26691	10948	76070	13663	9226	3670	26559	0.258787
35314	1024.5	16937	10774	3841	31552	5960	3717	1542	11219	0.262304
	829.5	21607	9172	2646	33425	7625	2842	1001	11468	0.255452
	1171.7	26622	16641	6921	50184	8748	5832	2178	16758	0.250336

Supplementary material B: Milk labelling pilot

Spreadsheet where the peptide intensities of the top 20 pup muscle proteins were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

		Light peptide intensity				Heavy peptide intensity				
Sample ID	Protein/peptide	1	2	3	Sum	1	2	3	Sum	IH/(IH+IL)
	P32848									
35313	1536.8	2183	1780	1039	5002	1778	774	606	3158	0.38701
	1420.7	112591	97000	47271	256862	49179	38998	19757	107934	0.295875
	788.5	8073	9598	4526	22197	7046	6096	2433	15575	0.412342
35314	1536.8	333767	286606	138827	759200	145443	117553	55081	318077	0.29526
	1420.7	302079	275252	132734	710065	137181	108244	52332	297757	0.295446
	788.5	236916	106050	27334	370300	103496	37416	8854	149766	0.287975
	P17751									
35313	1602.9	48369	46605	23620	118594	13567	12861	6704	33132	0.218367
	1466.7	90470	73659	36748	200877	24503	21016	9703	55222	0.215628
	758.4	113157	62246	20694	196097	37397	20051	5706	63154	0.243602
35314	1602.9	112059	104967	51897	268923	32984	30810	15446	79240	0.227595
	1466.7	195880	151243	80120	427243	56993	45604	22835	125432	0.226954
	758.4	286074	158119	52180	496373	94980	51857	18780	165617	0.250181
	P02089									
35313	1294.6	135785	99450	48634	283869	30417	22722	19197	72336	0.203074
	912.5	66538	35259	10311	112108	14661	7008	1723	23392	0.172635
	1091.6	126548	82167	30365	239080	22409	12509	4559	39477	0.14172
35314	1294.6	384733	284979	118978	788690	71568	51601	13221	136390	0.147436
	912.5	149516	79055	25239	253810	28655	14812	3256	46723	0.155467
	1091.6	187140	104435	40614	332189	31408	17173	6344	54925	0.141883
	P07310									
35313	1507.8	85231	76355	38362	199948	27073	21582	11143	59798	0.230217
	1723.9	249854	256101	142371	648326	134246	130280	72282	336808	0.341891
	1302.6	79253	54367	30841	164461	27116	16656	9733	53505	0.245474
35314	1507.8	263432	214438	106671	584541	87166	68448	32171	187785	0.243142

	1723.9	649448	639989	347492	1636929	361682	355454	197248	914384	0.358397
	1302.6	166329	117225	58881	342435	52985	37270	19547	109802	0.242797
	P21550									
35313	2743.3	21598	33553	28138	83289	8276	8703	8342	25321	0.233137
	1134.6	142568	105954	42536	291058	41702	28516	10973	81191	0.218109
	917.5	79966	36723	14982	131671	22667	10612	6644	39923	0.23266
35314	2743.3	64953	96472	86038	247463	22822	30342	23474	76638	0.236463
	1134.6	271450	177364	75513	524327	83545	54621	20383	158549	0.232178
	917.5	117185	56886	21472	195543	34259	16094	4967	55320	0.220519
	P05064									
35313	1342.7	112194	89890	37036	239120	38896	25671	12886	77453	0.244661
	951.5	143203	88220	31055	262478	54233	30732	10912	95877	0.267548
	1288.7	5999	4473	1843	12315	2093	1649	1000	4742	0.278009
35314	1342.7	100945	82060	35718	218723	34788	26533	12619	73940	0.252646
	951.5	224078	123569	47645	395292	82073	42878	14772	139723	0.261157
	1288.7	10453	7229	4056	21738	4382	3410	1902	9694	0.308412
	E9Q5U3									
35313	829.4	59978	27310	9116	96404	19494	8002	3149	30645	0.241206
	869.5	12347	8298	4040	24685	6046	2674	2559	11279	0.313619
	739.4	28484	12556	4980	46020	9357	4380	1970	15707	0.254459
35314	829.4	114782	51813	19108	185703	39814	17082	5757	62653	0.252271
	869.5	29890	17081	6800	53771	11380	4930	2686	18996	0.261052
	739.4	82961	39805	11357	134123	26988	10583	4447	42018	0.238548
	P16065									
35313	1361.7	10658	9911	4421	24990	3302	3490	916	7708	0.235733
	1942.9	48894	54865	33086	136845	15074	12687	8948	36709	0.211513
	1338.7	45116	34431	15941	95488	13703	9589	4622	27914	0.226204
35314	1361.7	37484	29102	14110	80696	10386	6781	3172	20339	0.201306
	1942.9	78047	82305	51281	211633	24155	17597	11529	53281	0.201126

	1338.7	89352	65454	27864	182670	21205	17286	6532	45023	0.197736
	Q9R0Y5									
35313	1130.7	41432	31646	11261	84339	13596	8832	3010	25438	0.231724
	2209.1	12626	18369	12368	43363	3745	5281	3283	12309	0.221099
	791.4	8900	4254	1710	14864	3288	1106	915	5309	0.263174
35314	1130.7	77872	52076	17666	147614	25897	14027	5155	45079	0.233942
	2209.1	48768	63934	49128	161830	18977	24797	15430	59204	0.26785
	791.4	19489	8016	3918	31423	6722	2697	1983	11402	0.266246
	P68134									
35313	1198.5	57285	40274	17617	115176	17520	10304	3970	31794	0.21633
	923.6	17848	9925	3907	31680	5092	2659	876	8627	0.214032
	1161.6	37861	25678	11326	74865	11820	6296	3407	21523	0.223295
35314	1198.5	107808	68983	29841	206632	32298	21126	9025	62449	0.232083
	923.6	48590	26641	9596	84827	14749	7714	3013	25476	0.230964
	1161.6	70621	41493	21627	133741	22297	12569	6963	41829	0.238247
	P11499									
35313	1242.7	15635	11693	4186	31514	5502	3730	1372	10604	0.251769
	1416.6	15151	12661	6173	33985	5383	4320	2669	12372	0.266885
	891.4	18811	9736	3673	32220	7490	4287	1802	13579	0.296491
35314	1242.7	27128	17733	7764	52625	10637	5791	2264	18692	0.262097
	1416.6	29806	22254	11150	63210	9358	8302	4089	21749	0.255994
	891.4	35723	20074	7120	62917	14587	6685	2452	23724	0.27382
	P09411									
35313	1219.7	31484	20717	8264	60465	9913	6754	3506	20173	0.250167
	1318.7	30544	26634	11139	68317	9592	7474	2895	19961	0.226115
	1769	27380	26815	17978	72173	8677	7885	5131	21693	0.231106
35314	1219.7	40138	25401	11771	77310	13490	8123	3501	25114	0.245196
	1318.7	46195	35488	17293	98976	13625	11413	4227	29265	0.228203
	1769	73003	69885	43206	186094	22679	22954	14431	60064	0.244006

	P52480									
35313	1636.9	30144	26864	14483	71491	9371	9375	4801	23547	0.247764
	1779.9	8815	9701	4667	23183	3013	2851	1880	7744	0.250396
	990.5	14541	8987	4091	27619	6129	3359	2257	11745	0.298369
35314	1636.9	61329	55007	31602	147938	20915	18155	9355	48425	0.24661
	1779.9	6167	6081	3714	15962	2404	1845	1196	5445	0.254356
	990.5	28237	14247	5943	48427	9009	4549	1626	15184	0.238701
	Q6PHC1									
35313	815.5	6425	3186	901	10512	1741	694	416	2851	0.21335
	800.3	85139	41395	11180	137714	25100	11104	2579	38783	0.219737
	2194.1	16393	24035	15354	55782	4983	5765	4984	15732	0.219985
35314	815.5	14447	6702	2246	23395	4559	1998	786	7343	0.23889
	800.3	194176	93003	26572	313751	65136	24194	7426	96756	0.235699
	2194.1	37570	51180	34975	123725	13141	12863	10087	36091	0.225828
	Q921I1									
35313	915.5	16449	9522	6181	32152	10781	5530	3614	19925	0.382607
	1238.6	21273	14909	6829	43011	11650	8313	3500	23463	0.352965
	1313.7	19996	14967	7554	42517	9873	7936	4361	22170	0.342727
35314	915.5	22259	12324	6053	40636	12375	7108	4487	23970	0.371018
	1238.6	53653	36443	17712	107808	26866	19503	8001	54370	0.335249
	1313.7	33586	27139	14379	75104	19221	14236	6665	40122	0.348203
	P62631									
35313	870.5	20419	9472	3221	33112	5518	2585	1001	9104	0.215653
	1120.6	20076	12384	4711	37171	5380	4293	1216	10889	0.226571
	1257.5	2457	1415	596	4468	954	478	311	1743	0.280631
35314	870.5	63817	30249	9493	103559	21016	9202	3134	33352	0.243604
	1120.6	26662	16877	6168	49707	9567	5631	1768	16966	0.254466
	1257.5	5926	3718	2306	11950	2224	1364	860	4448	0.271253
	Q5SX49									

35313	1616.9	7705	7015	3305	18025	2341	2334	864	5539	0.235062
	1213.6	21531	17793	6900	46224	7682	4606	2699	14987	0.244842
	1725.9	7083	7786	3763	18632	2591	3055	1148	6794	0.267207
35314	1616.9	17173	16650	9025	42848	5725	4629	2630	12984	0.232555
	1213.6	37192	28047	11220	76459	12595	7737	3705	24037	0.239184
	1725.9	11096	13849	8175	33120	2764	3754	2400	8918	0.212141
	P14152									
35313	917.5	24195	10937	4920	40052	6314	4320	1767	12401	0.236421
	1650.9	20875	20743	12487	54105	6153	6397	3491	16041	0.22868
	1007.5	18402	11271	5571	35244	6422	3058	1455	10935	0.236796
35314	917.5	42688	21005	7273	70966	14162	6825	3034	24021	0.252887
	1650.9	51432	49539	26443	127414	18563	15728	8683	42974	0.252213
	1007.5	345477	296701	149808	791986	134399	112199	52055	298653	0.273833
	O70250									
35313	1112.5	19618	13033	4705	37356	7234	4574	1523	13331	0.263006
	975.5	23045	14297	7158	44500	8407	4570	3252	16229	0.267236
	1360.6	22643	17796	9218	49657	6871	6582	4210	17663	0.262374
35314	1112.5	125180	78806	24733	228719	40937	22459	8541	71937	0.239267
	975.5	44942	25117	14438	84497	15516	9900	4854	30270	0.263752
	1360.6	32606	27247	15206	75059	11061	9116	7000	27177	0.265826
	P31001									
35313	1046.5	20177	11123	5426	36726	8115	4204	2003	14322	0.280559
	1268.6	21738	17519	8030	47287	10042	7961	3028	21031	0.30784
	1981	14907	18879	10989	44775	6088	7866	4066	18020	0.286966
35314	1046.5	38104	24253	9208	71565	16905	9747	4458	31110	0.302995
	1268.6	44269	39922	24789	108980	15014	11971	8759	35744	0.24698
	1981	26182	29779	20973	76934	8970	10705	6479	26154	0.253706

Supplementary material B: Milk labelling pilot

Spreadsheet where the peptide intensities of the low turnover liver protein were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

			Light peptide intensity				Heavy peptide intensity				
Sample ID	Day of labelling	Peptide	1	2	3	Sum	1	2	3	Sum	IH/(IH+IL)
35270	1	1583.8	7831	6537	3116	17484	263	158	115	536	0.029745
		1125.5	29499	15307	5720	50526	918	641	539	2098	0.039868
35271	1	1583.8	7505	5762	2603	15870	433	309	242	984	0.058384
		1125.5	68691	38339	15918	122948	3402	2007	1399	6808	0.052468
35280	2	1583.8	10801	9844	5181	25826	645	507	480	1632	0.059436
		1125.5	41244	22911	7994	72149	2209	1528	771	4508	0.058807
35288	3	1583.8	14962	13138	7717	35817	1945	1846	878	4669	0.115324
		1125.5	43041	23709	11101	77851	6538	3888	1784	12210	0.135575
35289	3	1583.8	15211	13593	6664	35468	2027	1932	1282	5241	0.128743
		1125.5	51294	31381	11494	94169	7526	4017	2102	13645	0.126561
35297	4	1583.8	12355	11337	5790	29482	2621	2547	1075	6243	0.174752
		1125.5	53417	32257	12633	98307	12445	6515	3298	22258	0.184614
35298	4	1583.8	24206	21077	11890	57173	4570	3787	1847	10204	0.151446
		1125.5	57175	33173	13896	104244	12172	6248	2686	21106	0.168377
35305	5	1583.8	12792	12720	6337	31849	3324	2146	1209	6679	0.173354
		1125.5	36094	21233	8020	65347	8492	6221	3382	18095	0.216857
35306	5	1583.8	20436	18006	9070	47512	4149	3359	1975	9483	0.166383
		1125.5	57448	32017	12425	101890	12566	7588	3757	23911	0.19007
35313	6	1583.8	26109	24605	15307	66021	9409	6424	4008	19841	0.23108
		1125.5	57674	35818	13354	106846	18158	10341	5031	33530	0.238858
35314	6	1583.8	16955	13691	6783	37429	4929	4300	2250	11479	0.234706
		1125.5	35913	20282	8528	64723	11106	6314	3374	20794	0.243156

Supplementary material B: Milk labelling pilot

Spreadsheet where the peptide intensities of the high turnover liver protein were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

Sample ID	Day of labelling	Peptide	1	2	3	Sum	1	2	3	Sum	IH/(IH+IL)
35270	1	916.5	14044	7598	2251	23893	515	577	285	1377	0.054491
		1046.5	11904	7213	3035	22152	829	818	335	1982	0.082125
35271	1	916.5	28552	14644	4960	48156	1541	1678	997	4216	0.080501
		1046.5	22134	13840	9001	44975	1281	2004	1486	4771	0.095907
35280	2	916.5	14538	7607	2439	24584	1785	1541	580	3906	0.137101
		1046.5	12085	7677	2481	22243	1298	875	1192	3365	0.131404
35288	3	916.5	26925	13836	4675	45436	4975	2616	1929	9520	0.173229
		1046.5	22970	13163	5285	41418	4517	2895	2205	9617	0.188439
35289	3	916.5	24541	16084	4219	44844	5481	3037	1391	9909	0.180976
		1046.5	22660	14595	4602	41857	4782	2850	2445	10077	0.194035
35297	4	916.5	27504	14800	5907	48211	8586	5012	2554	16152	0.250952
		1046.5	13954	9536	3958	27448	5183	2848	2970	11001	0.286119
35298	4	916.5	26303	13540	6080	45923	7938	4644	2065	14647	0.241819
		1046.5	23499	14770	5528	43797	7970	3586	5330	16886	0.278266
35305	5	916.5	12690	6716	2201	21607	4261	2759	2114	9134	0.297128
		1046.5	11757	8259	3636	23652	4587	2720	2404	9711	0.291071
35306	5	916.5	25424	13853	5255	44532	9251	6144	2202	17597	0.283233
		1046.5	12328	9805	4568	26701	4755	3641	2738	11134	0.294278
35313	6	916.5	21962	11742	4068	37772	10501	5523	2383	18407	0.327649
		1046.5	23808	15240	6535	45583	13019	7350	8818	29187	0.390357
35314	6	916.5	10316	6591	2770	19677	5070	3061	2813	10944	0.357402
		1046.5	5859	3032	1218	10109	2371	1450	2472	6293	0.383673

Supplementary material B: Milk labelling pilot

Spreadsheet where the peptide intensities of the low turnover muscle protein were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

Sample ID	Day of labelling	Peptide	1	2	3	Sum	1	2	3	Sum	IH/(IH+IL)
35270	1	800.4	24692	12883	2701	40276	1037	466	223	1726	0.041093
		1296.6	6692	3955	1898	12545	171	94	292	557	0.042513
35271	1	800.4	65067	31012	8226	104305	2036	1026	466	3528	0.032717
		1296.6	18919	14693	4961	38573	1042	1010	400	2452	0.059768
35280	2	800.4	43481	20909	5096	69486	2346	797	254	3397	0.046609
		1296.6	11778	9045	3901	24724	772	674	545	1991	0.074527
35288	3	800.4	110047	53910	14237	178194	12222	6279	1813	20314	0.102333
		1296.6	27086	20663	9300	57049	3631	3303	1648	8582	0.130761
35289	3	800.4	39287	18368	5847	63502	4789	2601	531	7921	0.110903
		1296.6	6509	5046	2966	14521	652	507	232	1391	0.087418
35297	4	800.4	78772	38196	11062	128030	15838	6859	1810	24507	0.160663
		1296.6	21175	13831	6979	41985	3531	2324	1439	7294	0.148014
35298	4	800.4	104455	51978	13953	170386	20184	8128	1621	29933	0.149427
		1296.6	26739	20139	10191	57069	5137	4217	2042	11396	0.16645
35305	5	800.4	127248	62198	18342	207788	33154	12007	3848	49009	0.190847
		1296.6	31681	23156	10575	65412	7553	6674	2793	17020	0.206473
35306	5	800.4	86814	44204	12033	143051	21061	8073	2165	31299	0.179518
		1296.6	21867	14979	7158	44004	4299	3114	1596	9009	0.169939
35313	6	800.4	103624	49747	14334	167705	31179	13411	3392	47982	0.222461
		1296.6	21288	15779	8071	45138	6407	5275	2982	14664	0.245209
35314	6	800.4	190130	91047	25789	306966	64132	23354	7229	94715	0.235797
		1296.6	55560	41008	23117	119685	17539	11004	6031	34574	0.22413

Supplementary material B: Milk labelling pilot

Spreadsheet where the peptide intensities of the high turnover muscle protein were recorded, along with the intensities of the corresponding labelled peptides, for the pups whose mothers were fed [$^{13}C_6$] lysine labelled diets over the course of six days. Precursor RIA values were calculated from this data.

			Light peptide intensity				Heavy peptide intensity				
Sample ID	Day of labelling	Peptide	1	2	3	Sum	1	2	3	Sum	IH/(IH+IL)
35270	1	1536.8	2990	1842	1006	5838	341	119	42	502	0.07918
		871.5	0	0	0	0	0	0	0	0	0
35271	1	1536.8	22580	18506	8227	49313	1325	964	452	2741	0.052657
		871.5	20532	9252	3121	32905	1270	579	752	2601	0.073255
35280	2	1536.8	17355	14630	7311	39296	1389	1299	606	3294	0.077342
		871.5	13250	8447	2242	23939	1518	548	298	2364	0.089876
35288	3	1536.8	24143	19587	9450	53180	5160	4028	1533	10721	0.167775
		871.5	39236	19980	9104	68320	9200	3784	2288	15272	0.182697
35289	3	1536.8	5777	4846	2556	13179	1452	1110	483	3045	0.187685
		871.5	10774	4929	1757	17460	2038	1692	412	4142	0.191742
35297	4	1536.8	69441	62969	31269	163679	19843	17120	7923	44886	0.215213
		871.5	22964	11289	3030	37283	6444	2512	713	9669	0.205934
35298	4	1536.8	55244	45047	23659	123950	15674	12934	6140	34748	0.218957
		871.5	27954	16483	5520	49957	9629	4360	1699	15688	0.238982
35305	5	1536.8	124119	106371	53539	284029	46547	34450	17214	98211	0.256935
		871.5	79183	42375	13284	134842	30846	14816	3896	49558	0.268753
35306	5	1536.8	93685	83724	37848	215257	33673	26772	12662	73107	0.253523
		871.5	64958	35889	11712	112559	22135	10348	2974	35457	0.239548
35313	6	1536.8	91724	84531	40084	216339	41126	32967	16274	90367	0.294637
		871.5	31172	17488	5144	53804	14700	6833	1802	23335	0.302506
35314	6	1536.8	377772	327584	158155	863511	165233	135351	64502	365086	0.297157
		871.5	222060	122405	33944	378409	99738	45451	14078	159267	0.296214

Supplementary material B: Communal nursing study

Spreadsheet where the peptide intensities of the low turnover pup liver and muscle proteins and the mother's urine were recorded, along with the intensities of the corresponding labelled peptides, where mothered were fed either a [D_4] or [D_9] lysine labelled diet over the course of the experiment. Precursor RIA values were calculated from this data. Pup IDs/mother IDs in red are pups from the [D_9] lysine labelled mother/mothers fed [D_9] lysine labelled diet. Those in green are the pups from the [D_4] lysine labelled mother/mothers fed [D_4] lysine labelled diet.

ID:	Sample type	Peptide	Light peptide				D4 peptide				D9 peptide				IH/(IH+IL)	
Unrelated1			1	2	3	Sum	1	2	3	Sum	1	2	3	Sum	D4	D9
36286	Liver	1584	62659	50278	23812	136749	9248	6102	3036	18386	9525	6562	3958	20045	0.1185	0.1278
	Muscle	800.4	70789	35054	8100	113943	9519	3770	1629	14918	11243	6050	1869	19162	0.1158	0.144
36287	Liver	1584	32757	24122	12609	69488	5754	3409	1976	11139	5157	3758	1973	10888	0.1382	0.1355
	Muscle	800.4	66521	30601	8912	106034	7591	4202	1080	12873	8680	4106	1288	14074	0.1083	0.1172
36288	Liver	1584	33920	25050	12844	71814	4984	3728	2103	10815	5404	3887	2047	11338	0.1309	0.1364
	Muscle	800.4	55375	25721	6902	87998	7855	3268	1335	12458	7347	3072	1279	11698	0.124	0.1173
36289	Liver	1584	69475	53292	22148	144915	10218	6953	3628	20799	12065	8916	3914	24895	0.1255	0.1466
	Muscle	800.4	53442	26612	6752	86806	6118	3184	720	10022	8536	3699	624	12859	0.1035	0.129
36290	Liver	1584	42839	30751	13309	86899	6577	4071	1702	12350	6175	4176	1890	12241	0.1244	0.1235
	Muscle	800.4	56296	26868	7518	90682	6246	3387	702	10335	7358	3380	1419	12157	0.1023	0.1182
36291	Liver	1584	70460	53261	23672	147393	10213	7714	3340	21267	11118	7279	3721	22118	0.1261	0.1305
	Muscle	800.4	105328	45943	11419	162690	13248	6049	1485	20782	13122	6034	2077	21233	0.1133	0.1154
36292	Liver	1584	30956	23555	10796	65307	5043	3983	1853	10879	4084	3415	1371	8870	0.1428	0.1196
	Muscle	800.4	67357	30044	8663	106064	8951	4134	1269	14354	7292	3210	1400	11902	0.1192	0.1009

36293	Liver	1584	61972	45058	18539	125569	8512	6497	2700	17709	9896	7755	3477	21128	0.1236	0.144
	Muscle	800.4	80072	36395	10212	126679	9237	4571	1000	14808	12498	4595	1504	18597	0.1047	0.128
36294	Liver	1584	36036	25617	10978	72631	6081	4300	1602	11983	5107	4007	1911	11025	0.1416	0.1318
	Muscle	800.4	22327	10440	2914	35681	3475	1576	336	5387	2880	1336	346	4562	0.1312	0.1134
36295	Liver	1584	72826	49803	20547	143176	12429	9079	3467	24975	12904	7635	4448	24987	0.1485	0.1486
	Muscle	800.4	41940	20567	5772	68279	6112	2613	874	9599	6043	2367	900	9310	0.1233	0.12
36296	Urine		66721	54513	30412	151646	66595	54619	30767	151981	2350	1099	265	3714	0.5006	0.0239
36297	Urine		14018	13025	7933	34976	1991	555	0	2546	14705	15612	8523	38840	0.0679	0.5262
Related 1																
36329	Liver	1584	62558	46735	20572	129865	6963	4882	2353	14198	6233	4982	1834	13049	0.0986	0.0913
	Muscle	800.4	251735	117211	32495	401441	24854	11797	3808	40459	30150	12475	4728	47353	0.0916	0.1055
36330	Liver	1584	44790	32285	14495	91570	5286	3655	1511	10452	4135	3806	1400	9341	0.1024	0.0926
	Muscle	800.4	77240	34437	8933	120610	8182	3379	1185	12746	8019	3533	1218	12770	0.0956	0.0957
36331	Liver	1584	30444	14101	5527	50072	2524	1133	419	4076	3403	1341	634	5378	0.0753	0.097
	Muscle	800.4	211721	95185	24988	331894	17387	7885	2390	27662	22114	9396	3272	34782	0.0769	0.0949

36332	Liver	1584	30762	21729	9839	62330	3898	2477	1189	7564	2543	1645	891	5079	0.1082	0.0753
	Muscle	800.4	144699	63893	16896	225488	12444	6633	1958	21035	11716	5554	2194	19464	0.0853	0.0795
36333	Liver	1584	42929	29618	12687	85234	5396	3557	1690	10643	3944	3002	1348	8294	0.111	0.0887
	Muscle	800.4	96984	43402	11443	151829	8876	4352	1284	14512	8259	4432	1007	13698	0.0872	0.0828
36334	Liver	1584	16500	11289	5296	33085	2184	1624	750	4558	1427	1089	584	3100	0.1211	0.0857
	Muscle	800.4	105716	45361	13193	164270	10641	5732	1335	17708	10498	5059	1500	17057	0.0973	0.0941
36335	Liver	1584	37731	28152	12362	78245	4792	2997	1088	8877	4143	2961	1363	8467	0.1019	0.0976
	Muscle	800.4	144335	62950	18345	225630	14249	6557	1640	22446	14921	7352	2720	24993	0.0905	0.0997
36336	Liver	1584	22466	17037	6460	45963	3000	1649	920	5569	2443	1531	1093	5067	0.1081	0.0993
	Muscle	800.4	167813	74511	19238	261562	16534	7460	2422	26416	19173	8895	2663	30731	0.0917	0.1051
36337	Liver	1584	27721	19873	8262	55856	3694	2266	1036	6996	2533	1423	927	4883	0.1113	0.0804
	Muscle	800.4	106998	45407	13414	165819	11128	5765	1519	18412	9260	4034	1575	14869	0.0999	0.0823
36338	Liver	1584	13000	8364	3661	25025	1479	1258	619	3356	1276	912	608	2796	0.1182	0.1005
	Muscle	800.4	99434	47737	13469	160640	10251	4329	1508	16088	11090	4525	2115	17730	0.091	0.0994
36339	Urine		73102	65273	34438	172813	72721	62716	30898	166335	4339	1850	1079	7268	0.4904	0.0404
36340	Urine		80169	71733	36479	188381	4968	2788	741	8497	30209	24939	12455	67603	0.0432	0.2641

Related 2																
36472	Liver	1584	35893	25869	12406	74168	6376	4820	2381	13577	3531	1977	1263	6771	0.1547	0.0837
	Muscle	800.4	77329	34669	10049	122047	11689	5954	2009	19652	6473	2883	1753	11109	0.1387	0.0834
36473	Liver	1584	54330	38992	18668	111990	8587	5713	2249	16549	3666	2971	1253	7890	0.1287	0.0658
	Muscle	800.4	83234	39286	12174	134694	13742	10294	3785	27821	10585	4806	2432	17823	0.1712	0.1169
36474	Liver	1584	30824	23940	11093	65857	6684	5649	2464	14797	3639	1929	1175	6743	0.1835	0.0929
	Muscle	800.4	62180	29570	6820	98570	11733	5633	1462	18828	5851	2302	996	9149	0.1604	0.0849
36475	Liver	1584	38044	30249	13171	81464	7473	5017	2565	15055	2693	2151	989	5833	0.156	0.0668
	Muscle	800.4	17499	8634	2771	28904	3294	1759	627	5680	1941	921	732	3594	0.1642	0.1106
36476	Urine		22791	22976	11566	57333	27313	22802	11669	61784	2003	1652	764	4419	0.5187	0.0716
36477	Urine		1701	1857	897	4455	257	171	103	531	1149	936	586	2671	0.1065	0.3748
Unrelated2																
36679	Liver	1584	24123	14967	5366	44456	2609	2209	838	5656	3493	2093	930	6516	0.1129	0.1278
	Muscle	800.4	24329	11133	2990	38452	2440	1106	281	3827	2768	2008	371	5147	0.0905	0.1181
36680	Liver	1584	18760	10440	4470	33670	2569	1671	764	5004	2659	1798	769	5226	0.1294	0.1344

	Muscle	800.4	46399	21059	5663	73121	5759	2405	916	9080	7060	3316	1135	11511	0.1105	0.136
36681	Liver	1584	15625	10186	4592	30403	2588	1872	789	5249	2989	2010	724	5723	0.1472	0.1584
	Muscle	800.4	38116	18724	4677	61517	5861	2879	822	9562	6893	3274	913	11080	0.1345	0.1526
36682	Liver	1584	19886	15560	7921	43367	4215	3105	1364	8684	2801	2265	1796	6862	0.1668	0.1366
	Muscle	800.4	26479	13186	3755	43420	2994	1601	566	5161	3340	2287	741	6368	0.1062	0.1279
36683	Liver	1584	21828	13591	5574	40993	2842	2136	798	5776	2176	1965	1100	5241	0.1235	0.1134
	Muscle	800.4	118734	53806	14781	187321	17235	7782	2428	27445	17050	7972	2727	27749	0.1278	0.129
36684	Liver	1584	9640	6730	2781	19151	1826	858	484	3168	1799	1046	511	3356	0.1419	0.1491
	Muscle	800.4	18064	10401	2395	30860	2038	1299	348	3685	3317	1567	315	5199	0.1067	0.1442
36685	Liver	1584	19920	10248	4700	34868	2547	1338	607	4492	2818	1537	662	5017	0.1141	0.1258
	Muscle	800.4	24122	10783	2916	37821	2908	1325	393	4626	3090	1456	416	4962	0.109	0.116
36686	Liver	1584	24778	14039	6583	45400	3098	2119	874	6091	3939	2618	995	7552	0.1183	0.1426
	Muscle	800.4	21048	10089	3091	34228	2219	1393	367	3979	3410	1690	442	5542	0.1041	0.1394
36687	Liver	1584	3232	1717	757	5706	596	378	299	1273	526	283	210	1019	0.1824	0.1515
	Muscle	800.4	62463	28981	8628	100072	7106	3520	1169	11795	9498	4235	1549	15282	0.1054	0.1325
36478	Urine		82028	79470	46306	207804	13508	7976	4257	25741	83474	78744	44778	206996	0.1102	0.499

36688	Urine		16890	14329	8559	39778	19829	18432	9852	48113	711	249	177	1137	0.5474	0.0278
Related 3																
37251	Liver	1584	8116	4338	2341	14795	946	673	414	2033	995	510	315	1820	0.1208	0.1095
	Muscle	800.4	913	408	217	1538	142	32	0	174	149	36	19	204	0.1016	0.1171
37252	Liver	1584	21668	15067	6112	42847	2557	1297	402	4256	2993	1927	620	5540	0.0904	0.1145
	Muscle	800.4	73388	33982	9935	117305	7347	3941	1131	12419	9224	4726	1778	15728	0.0957	0.1182
37253	Liver	1584	29461	20421	9797	59679	3295	2118	1033	6446	3162	2521	1404	7087	0.0975	0.1061
	Muscle	800.4	97911	43096	13286	154293	8672	4223	1225	14120	12310	5133	2067	19510	0.0838	0.1123
37254	Liver	1584	60316	49460	21085	130861	7464	5233	1847	14544	8489	5944	2769	17202	0.1	0.1162
	Muscle	800.4	104543	46913	13029	164485	10244	5276	1792	17312	17029	7658	3483	28170	0.0952	0.1462
37255	Liver	1584	23620	16765	7961	48346	2656	1384	650	4690	2899	2313	958	6170	0.0884	0.1132
	Muscle	800.4	116539	54332	14794	185665	9241	5587	1459	16287	15881	6974	2647	25502	0.0806	0.1208
37256	Liver	1584	20979	13141	6560	40680	1823	1002	572	3397	2943	1399	836	5178	0.0771	0.1129
	Muscle	800.4	37119	16849	3361	57329	3239	1540	313	5092	5548	2038	834	8420	0.0816	0.1281
37257	Liver	1584	18861	13284	5741	37886	1801	1260	531	3592	1335	1345	845	3525	0.0866	0.0851

	Muscle	800.4	18238	8947	2816	30001	2191	679	278	3148	1570	1334	441	3345	0.095	0.1003
37258	Urine		118404	10639	62920	287723	8478	2797	978	12253	109114	98158	54222	261494	0.0408	0.4761
37259	Urine		158293	143693	74462	376448	149868	123494	64412	337774	3604	2295	880	6779	0.4729	0.0177
Related 4																
37260	Liver	1584	33593	28667	11869	74129	5236	4477	1793	11506	5656	3428	2228	11312	0.1344	0.1324
	Muscle	800.4	42786	22714	5161	70661	5783	2443	780	9006	6661	2537	827	10025	0.113	0.1242
37261	Liver	1584	128969	98190	40326	267485	15888	10376	5792	32056	19578	13035	4958	37571	0.107	0.1232
	Muscle	800.4	34912	15124	4419	54455	3016	1257	610	4883	4834	1957	689	7480	0.0823	0.1208
37262	Liver	1584	19931	14031	5671	39633	2943	2437	824	6204	3949	2165	879	6993	0.1353	0.15
	Muscle	800.4	42689	18635	5307	66631	6057	3188	925	10170	7603	3461	1478	12542	0.1324	0.1584
37263	Liver	1584	24284	13675	5474	43433	2399	1258	702	4359	3303	1949	1146	6398	0.0912	0.1284
	Muscle	800.4	46530	21236	6599	74365	4800	2011	827	7638	8213	3563	922	12698	0.0931	0.1458
37264	Liver	1584	7101	4811	2366	14278	1380	729	367	2476	944	835	558	2337	0.1478	0.1407
	Muscle	800.4	35429	17009	4120	56558	5396	2535	586	8517	6902	4205	1117	12224	0.1309	0.1777
37265	Liver	1584	5234	5382	2678	13294	984	907	328	2219	973	610	294	1877	0.143	0.1237

	Muscle	800.4	32401	16756	4713	53870	3585	1542	533	5660	4169	2462	783	7414	0.0951	0.121
37266	Liver	1584	27094	20405	7899	55398	4406	2964	1163	8533	4444	2390	1112	7946	0.1335	0.1254
	Muscle	800.4	38485	18023	5228	61736	6197	2560	748	9505	6020	3428	1430	10878	0.1334	0.1498
37267	Liver	1584	9272	8148	3106	20526	1207	910	798	2915	1004	933	538	2475	0.1244	0.1076
	Muscle	800.4	29459	12703	3540	45702	3094	1540	467	5101	3111	2031	337	5479	0.1004	0.1071
37268	Liver	1584	32232	19876	7902	60010	4260	2418	1253	7931	3988	2800	1118	7906	0.1167	0.1164
	Muscle	800.4	61255	31037	7757	100049	6434	3481	1207	11122	10504	4258	1919	16681	0.1	0.1429
37269	Liver	1584	12289	9987	3717	25993	2132	1428	657	4217	2152	1235	890	4277	0.1396	0.1413
	Muscle	800.4	27818	11879	3530	43227	2866	1570	433	4869	4370	1450	692	6512	0.1012	0.1309
37270	Liver	1584	140339	105222	48473	294034	17019	12290	5925	35234	20643	14813	6603	42059	0.107	0.1251
	Muscle	800.4	177721	73562	19388	270671	17780	7206	3269	28255	23753	10434	6559	40746	0.0945	0.1308
37271	Liver	1584	119011	82993	37379	239383	17209	10874	5546	33629	17046	12109	6044	35199	0.1232	0.1282
	Muscle	800.4	63977	26325	8836	99138	7408	2976	677	11061	9166	4408	3005	16579	0.1004	0.1433
37272	Liver	1584	174353	125030	52136	351519	21507	13329	6766	41602	24686	14861	7843	47390	0.1058	0.1188
	Muscle	800.4	86243	35108	8790	130141	7470	4009	911	12390	10345	4200	3180	17725	0.0869	0.1199
37273	Liver	1584	82384	59560	25579	167523	15255	8968	3799	28022	10581	8523	3737	22841	0.1433	0.12

	Muscle	800.4	63078	29385	7610	100073	8205	4098	885	13188	7773	3823	3126	14722	0.1164	0.1282
37274	Liver	1584	45649	31748	13156	90553	7123	5386	2690	15199	5633	3509	2119	11261	0.1437	0.1106
	Muscle	800.4	95287	40577	11018	146882	13338	5932	1612	20882	11643	5373	4867	21883	0.1245	0.1297
37275	Urine		64364	64272	36255	164891	5702	2093	665	8460	63194	63570	37081	163845	0.0488	0.4984
37276	Urine		44886	40171	21888	106945	52548	45668	26814	125030	1720	819	477	3016	0.539	0.0274
Related 5																
37293	Liver	1584	109536	76780	33426	219742	17632	11275	4713	33620	12141	7679	4212	24032	0.1327	0.0986
	Muscle	800.4	177188	72796	19803	269787	22643	10163	2698	35504	18023	8778	5833	32634	0.1163	0.1079
37294	Liver	1584	101804	78463	40614	220881	17648	13691	6255	37594	12626	8787	5054	26467	0.1454	0.107
	Muscle	800.4	92041	41573	11106	144720	10942	5418	1680	18040	10363	4228	1129	15720	0.1108	0.098
37295	Liver	1584	93288	77660	40750	211698	19537	16253	8172	43962	12930	10676	6663	30269	0.172	0.1251
	Muscle	800.4	134128	59423	15406	208957	21137	10003	2634	33774	13482	5928	1747	21157	0.1391	0.0919
37296	Liver	1584	77059	61692	27434	166185	14240	10370	3526	28136	8609	7800	4095	20504	0.1448	0.1098
	Muscle	800.4	169162	69483	20303	258948	18753	8747	1836	29336	17187	8338	1868	27393	0.1018	0.0957
37297	Liver	1584	100460	79515	33901	213876	19532	17046	7072	43650	13357	9438	4974	27769	0.1695	0.1149

	Muscle	800.4	137030	60074	16835	213939	22928	10258	2628	35814	18004	7953	2214	28171	0.1434	0.1164
37298	Liver	1584	75715	67004	29810	172529	16336	15948	6236	38520	11246	8942	4758	24946	0.1825	0.1263
	Muscle	800.4	114487	54152	12560	181199	16575	7555	2359	26489	11725	4754	1767	18246	0.1275	0.0915
37299	Liver	1584	72299	43159	17900	133358	10978	7506	2920	21404	10834	5556	2749	19139	0.1383	0.1255
	Muscle	800.4	125185	53309	14614	193108	16351	8774	2400	27525	15499	6447	1756	23702	0.1248	0.1093
37300	Liver	1584	116660	80006	34827	231493	24020	15860	6961	46841	13144	8896	3913	25953	0.1683	0.1008
	Muscle	800.4	68451	27694	7238	103383	10005	5615	1582	17202	7982	2485	713	11180	0.1427	0.0976
37301	Liver	1584	67496	61211	28980	157687	12972	10503	3017	26492	8107	5645	3091	16843	0.1438	0.0965
	Muscle	800.4	193809	76306	22725	292840	25155	11360	3626	40141	16303	6977	1782	25062	0.1206	0.0788
37302	Liver	1584	80799	48565	19318	148682	13333	9381	3905	26619	10664	7295	2677	20636	0.1518	0.1219
	Muscle	800.4	107945	45540	12412	165897	13609	6069	1918	21596	10816	5736	1164	17716	0.1152	0.0965
37303	Liver	1584	97070	76689	30596	204355	18873	13761	5604	38238	12501	8529	3421	24451	0.1576	0.1069
	Muscle	800.4	127288	52614	13164	193066	18307	7696	2311	28314	15904	6041	1706	23651	0.1279	0.1091
37304	Urine		26480	23866	12503	62849	29250	25488	14450	69188	758	335	0	1093	0.524	0.0171
37305	Urine		2065	2259	1119	5443	492	99	0	591	1031	1249	1090	3370	0.0979	0.3824
NC1																

37306	Liver	1584	31390	21589	10389	63368	0	0	0	0	4832	3257	1728	9817	0	0.1341
37307	Liver	1584	27986	19728	9203	56917	0	0	0	0	4404	3175	1349	8928	0	0.1356
37308	Liver	1584	22066	15492	7275	44833	0	0	0	0	3237	2558	1389	7184	0	0.1381
37309	Liver	1584	42216	30968	13725	86909	0	0	0	0	6055	4672	2182	12909	0	0.1293
37310	Liver	1584	6470	4376	2136	12982	0	0	0	0	1036	883	337	2256	0	0.1481
37311	Liver	1584	57744	40969	17724	116437	0	0	0	0	9655	6460	2871	18986	0	0.1402
37312	Liver	1584	44543	31377	13482	89402	0	0	0	0	6304	4057	2211	12572	0	0.1233
37313	Urine		52208	52022	29879	134109	4557	1306	922	6785	61415	57571	33339	152325	0.0482	0.5318
37314	Urine		213208	172937	78849	464994	48916	37813	18272	105001	8602	6340	2997	17939	0.1842	0.0371
Related 6																
37316	Liver	1584	16978	11427	4627	33032	2510	1368	703	4581	1904	1721	762	4387	0.1218	0.1172
	Muscle	800.4	67371	26953	8397	102721	7425	3286	1071	11782	8102	3660	1294	13056	0.1029	0.1128
37317	Liver	1584	45725	32564	13959	92248	5034	3512	1844	10390	7304	5170	2616	15090	0.1012	0.1406
	Muscle	800.4	162264	73592	19161	255017	13061	6569	1661	21291	23784	11208	3215	38207	0.0771	0.1303
37318	Liver	1584	60908	43835	19359	124102	7426	4967	2460	14853	7915	5969	2964	16848	0.1069	0.1195

	Muscle	800.4	169834	77430	21337	268601	16258	7324	2603	26185	20878	9509	3346	33733	0.0888	0.1116
37319	Liver	1584	25567	18977	7625	52169	3852	2158	1158	7168	3953	2454	1473	7880	0.1208	0.1312
	Muscle	800.4	84554	39239	10257	134050	9022	3734	1880	14636	10955	5073	1669	17697	0.0984	0.1166
37320	Liver	1584	43388	32680	16262	92330	4857	3575	1763	10195	7629	5323	2779	15731	0.0994	0.1456
	Muscle	800.4	89236	40986	10630	140852	7498	3329	1121	11948	12063	5483	2947	20493	0.0782	0.127
37321	Liver	1584	61525	43880	20736	126141	10104	6389	3042	19535	9420	6505	2999	18924	0.1341	0.1305
	Muscle	800.4	79644	35924	10658	126226	8914	4181	1262	14357	10526	3568	1651	15745	0.1021	0.1109
37322	Liver	1584	31779	23205	10598	65582	3225	2118	823	6166	4512	2940	1803	9255	0.0859	0.1237
	Muscle	800.4	125083	61608	15607	202298	9142	5197	1728	16067	17137	8407	2255	27799	0.0736	0.1208
37323	Liver	1584	44323	31885	14451	90659	5741	4402	1790	11933	7176	4868	2218	14262	0.1163	0.1359
	Muscle	800.4	71621	32351	8801	112773	6039	3356	1118	10513	9047	4455	1864	15366	0.0853	0.1199
37324	Liver	1584	24315	17499	8462	50276	2591	1934	796	5321	4240	2724	1194	8158	0.0957	0.1396
	Muscle	800.4	155055	73022	17766	245843	11659	6248	1456	19363	23197	11064	3554	37815	0.073	0.1333
37325	Liver	1584	36826	28157	12820	77803	3782	2871	1629	8282	7466	5289	2732	15487	0.0962	0.166
	Muscle	800.4	76848	33302	8156	118306	4851	3236	1324	9411	11720	4684	2737	19141	0.0737	0.1393
37326	Liver	1584	8317	5512	2609	16438	836	572	247	1655	1107	933	663	2703	0.0915	0.1412

	Muscle	800.4	139505	66975	18065	224545	14024	5725	1611	21360	20032	10186	3124	33342	0.0869	0.1293
37327	Liver	1584	26321	18797	8568	53686	4248	2312	1726	8286	3734	2477	1234	7445	0.1337	0.1218
	Muscle	800.4	50404	24860	6097	81361	6208	3020	725	9953	6123	3017	1058	10198	0.109	0.1114
37328	Liver	1584	37882	26574	13059	77515	5044	3325	1595	9964	6006	5056	2474	13536	0.1139	0.1487
	Muscle	800.4	118700	58007	13478	190185	11065	5266	1510	17841	16664	8384	2446	27494	0.0858	0.1263
37729	Urine		156813	139284	72987	369084	23785	14344	6574	44703	63095	54035	28908	146038	0.108	0.2835
37730	Urine		35837	32440	17934	86211	36411	30276	14088	80775	913	375	199	1487	0.4837	0.017
Unrelated3																
37731	Liver	1584	122038	93489	40491	256018	18687	13008	5907	37602	13830	9160	5571	28561	0.1281	0.1004
	Muscle	800.4	76792	34485	8517	119794	9306	5120	1357	15783	7978	3848	951	12777	0.1164	0.0964
37732	Liver	1584	34192	25758	12368	72318	5978	3846	1964	11788	3790	2848	1378	8016	0.1402	0.0998
	Muscle	800.4	97757	41606	10726	150089	12769	6190	2314	21273	11444	5049	1170	17663	0.1241	0.1053
37733	Liver	1584	31891	22324	9969	64184	4880	3210	1746	9836	6955	4130	2309	13394	0.1329	0.1727
	Muscle	800.4	29877	13783	3481	47141	3840	1750	890	6480	5286	2476	841	8603	0.1208	0.1543
37734	Liver	1584	43975	32630	15038	91643	6005	4331	2339	12675	8377	5595	2985	16957	0.1215	0.1561
	Muscle	800.4	42895	19400	5262	67557	4837	2216	628	7681	7063	3398	906	11367	0.1021	0.144

37735	Liver	1584	34952	24601	11383	70936	6854	3887	1556	12297	3729	2859	1347	7935	0.1477	0.1006
	Muscle	800.4	87331	38333	9869	135533	13966	6329	1982	22277	9139	4365	1326	14830	0.1412	0.0986
37736	Liver	1584	81483	58418	27138	167039	14710	10296	4974	29980	16410	10259	6637	33306	0.1522	0.1662
	Muscle	800.4	52748	22748	6460	81956	7996	3220	868	12084	10297	4365	1317	15979	0.1285	0.1632
37737	Liver	1584	31775	22044	9349	63168	5706	3570	1569	10845	5509	4174	1862	11545	0.1465	0.1545
	Muscle	800.4	22106	10185	3205	35496	3370	1887	647	5904	4728	2485	654	7867	0.1426	0.1814
37738	Liver	1584	35806	25599	12223	73628	5107	3890	1601	10598	6236	5456	2017	13709	0.1258	0.157
	Muscle	800.4	51499	21852	7192	80543	5984	2643	1180	9807	9697	4818	1506	16021	0.1085	0.1659
37739	Liver	1584	130300	90370	43495	264165	22626	16680	6780	46086	13736	9295	5498	28529	0.1485	0.0975
	Muscle	800.4	114510	48825	13410	176745	17368	7131	1844	26343	11010	4803	1639	17452	0.1297	0.0899
37740	Liver	1584	60804	44093	19986	124883	10406	6272	2616	19294	7141	5437	2234	14812	0.1338	0.106
	Muscle	800.4	84090	38589	11299	133978	10910	5434	1589	17933	8439	3809	1238	13486	0.118	0.0915
37741	Liver	1584	49033	37330	15389	101752	7883	5336	2216	15435	8956	6827	3469	19252	0.1317	0.1591
	Muscle	800.4	40440	16947	5188	62575	4356	2484	850	7690	6378	3321	1014	10713	0.1094	0.1462
37742	Liver	1584	53202	40052	18238	111492	8110	5202	3696	17008	5481	4034	2085	11600	0.1324	0.0942
	Muscle	800.4	217244	92211	27676	337131	26678	11376	3294	41348	21781	9225	2911	33917	0.1092	0.0914

37743	Liver	1584	40579	29149	12974	82702	6893	5078	2551	14522	3914	2664	1645	8223	0.1494	0.0904
	Muscle	800.4	38289	16673	4847	59809	4865	2215	652	7732	6849	3493	1048	11390	0.1145	0.16
37744	Liver	1584	33210	25864	11188	70262	5929	4685	1962	12576	3720	2620	1767	8107	0.1518	0.1034
	Muscle	800.4	73716	31898	9442	115056	7946	3781	954	12681	13672	6550	1565	21787	0.0993	0.1592
37745	Liver	1584	41730	32774	15339	89843	5768	3421	1291	10480	8353	6180	2575	17108	0.1045	0.16
	Muscle	800.4	106801	48199	13864	168864	13614	6428	2169	22211	9220	4713	1339	15272	0.1162	0.0829
37746	Urine		9356	7735	4869	21960	10670	8989	4966	24625	914	299	63	1276	0.5286	0.0549
37747	Urine		12958	13041	8239	34238	1461	466	133	2060	13828	10853	6782	31463	0.0568	0.4789
Unrelated4																
37748	Liver	1584	41228	27761	11991	80980	6126	4381	2353	12860	5674	3775	3026	12475	0.137	0.1335
	Muscle	800.4	12504	5743	1721	19968	1684	1258	437	3379	1870	912	396	3178	0.1447	0.1373
37749	Liver	1584	29839	21887	8659	60385	6296	4953	2012	13261	4673	3655	1930	10258	0.1801	0.1452
	Muscle	800.4	13318	6431	1833	21582	2720	1158	193	4071	1966	719	302	2987	0.1587	0.1216
37750	Liver	1584	35163	23229	10852	69244	6599	4352	2836	13787	4928	3451	1625	10004	0.166	0.1262
	Muscle	800.4	11228	5643	1362	18233	1568	1432	223	3223	1943	757	111	2811	0.1502	0.1336
37751	Liver	1584	10116	6834	2777	19727	2461	1485	905	4851	1277	818	356	2451	0.1974	0.1105

	Muscle	800.4	8503	3652	1237	13392	2358	723	139	3220	847	693	173	1713	0.1938	0.1134
37752	Liver	1584	19140	12948	6491	38579	4332	2599	1075	8006	2627	1434	1611	5672	0.1719	0.1282
	Muscle	800.4	18427	8227	2061	28715	2803	1015	661	4479	3181	927	367	4475	0.1349	0.1348
37622	Urine		72167	64017	33481	169665	70285	57302	29169	156756	2044	1722	336	4102	0.4802	0.0236
37623	Urine		48314	43186	24861	116361	4816	1993	659	7468	51572	45159	27240	123971	0.0603	0.5158
Unrelated5																
37753	Liver	1584	23590	15563	7816	46969	4387	2952	1494	8833	2831	1390	1224	5445	0.1583	0.1039
	Muscle	800.4	29252	12623	3245	45120	4958	2137	509	7604	3521	1296	641	5458	0.1442	0.1079
37754	Liver	1584	17295	11519	4751	33565	4376	2881	1060	8317	2364	1948	831	5143	0.1986	0.1329
	Muscle	800.4	7413	4060	764	12237	1342	867	145	2354	754	357	142	1253	0.1613	0.0929
37755	Liver	1584	12926	7284	3631	23841	2247	1560	687	4494	1519	1105	472	3096	0.1586	0.1149
	Muscle	800.4	5465	3161	952	9578	979	490	124	1593	784	180	115	1079	0.1426	0.1012
37756	Liver	1584	26970	20916	7911	55797	6040	4871	2076	12987	3997	3039	1822	8858	0.1888	0.137
	Muscle	800.4	14670	6717	1730	23117	2363	1489	407	4259	1624	1489	435	3548	0.1556	0.1331
37757	Liver	1584	30335	19133	8277	57745	7944	5648	2063	15655	3996	2459	1595	8050	0.2133	0.1223
	Muscle	800.4	28056	11518	3569	43143	4944	3231	693	8868	2196	1312	418	3926	0.1705	0.0834

37758	Liver	1584	10667	7463	3295	21425	3566	2067	1082	6715	1876	1915	645	4436	0.2386	0.1715
	Muscle	800.4	4491	2236	503	7230	1033	473	358	1864	509	558	206	1273	0.205	0.1497
37759	Liver	1584	45873	32507	13460	91840	9808	6889	3156	19853	4425	3697	966	9088	0.1777	0.09
	Muscle	800.4	24244	10308	2808	37360	5043	1910	568	7521	2420	923	325	3668	0.1676	0.0894
37760	Liver	1584	54665	38372	16250	109287	9802	6408	3734	19944	5387	3972	1811	11170	0.1543	0.0927
	Muscle	800.4	26233	11259	2604	40096	4782	2320	643	7745	2647	791	386	3824	0.1619	0.0871
37761	Liver	1584	18169	11809	5291	35269	4426	3080	948	8454	2322	1574	935	4831	0.1934	0.1205
	Muscle	800.4	14050	5813	2134	21997	2291	884	446	3621	937	699	139	1775	0.1413	0.0747
37762	Liver	1584	16365	11703	4726	32794	3098	2195	984	6277	1690	988	618	3296	0.1607	0.0913
	Muscle	800.4	42514	18052	6358	66924	6942	3756	1113	11811	4488	2055	736	7279	0.15	0.0981
37763	Urine		24093	19315	11404	54812	29777	23202	12746	65725	918	592	469	1979	0.5453	0.0348
37764	Urine		17475	18979	10386	46840	2234	826	236	3296	18339	19181	10627	48147	0.0657	0.5069
Related 7																
37765	Liver	1584	18111	12024	4955	35090	3624	2096	1227	6947	2109	1055	892	4056	0.1653	0.1036
	Muscle	800.4	48584	22160	5554	76298	6259	2810	1031	10100	3457	2019	610	6086	0.1169	0.0739

37766	Liver	1584	14472	8703	3774	26949	2705	1640	879	5224	1987	1442	512	3941	0.1624	0.1276
	Muscle	800.4	18751	7590	1938	28279	2612	1399	448	4459	1317	516	391	2224	0.1362	0.0729
37767	Liver	1584	30208	18298	9238	57744	5491	3730	1380	10601	2972	1852	1112	5936	0.1551	0.0932
	Muscle	800.4	22212	9861	3421	35494	3345	1596	839	5780	1421	651	244	2316	0.14	0.0613
37768	Liver	1584	26002	16340	6699	49041	4069	2904	1282	8255	2118	2200	1937	6255	0.1441	0.1131
	Muscle	800.4	26337	12593	3519	42449	4935	1578	439	6952	2239	1338	471	4048	0.1407	0.0871
37769	Liver	1584	32895	22531	8666	64092	5372	3978	1859	11209	3362	2105	1463	6930	0.1489	0.0976
	Muscle	800.4	33351	15394	4300	53045	4001	1792	650	6443	2699	1177	244	4120	0.1083	0.0721
37770	Liver	1584	33943	23649	9652	67244	6048	4103	1746	11897	3651	2335	1640	7626	0.1503	0.1019
	Muscle	800.4	51634	23625	5676	80935	7830	2822	615	11267	3736	1380	654	5770	0.1222	0.0665
37771	Liver	1584	13328	8730	2883	24941	2970	1668	1015	5653	1853	848	491	3192	0.1848	0.1135
	Muscle	800.4	13858	5609	1614	21081	1600	689	329	2618	608	350	315	1273	0.1105	0.0569
37772	Liver	1584	31209	22480	9997	63686	4731	3746	1887	10364	3203	2076	1377	6656	0.14	0.0946
	Muscle	800.4	24777	10940	2997	38714	2809	1246	463	4518	1275	404	239	1918	0.1045	0.0472
37773	Liver	1584	60976	41188	18941	121105	10603	6731	2941	20275	5134	3638	2752	11524	0.1434	0.0869
	Muscle	800.4	15151	7747	1787	24685	1605	617	309	2531	553	492	163	1208	0.093	0.0467

37774	Urine		51779	47342	27831	126952	4078	1500	565	6143	43624	41261	24875	109760	0.0462	0.4637
37775	Urine		89990	79417	42954	212361	109862	98747	50592	259201	2241	814	421	3476	0.5497	0.0161
NC2																
37776	Liver	1584	10110	6843	2891	19844	2643	1809	1003	5455	0	0	0	0	0.2156	0
	Muscle	800.4	137252	64578	19023	220853	34922	14181	4692	53795	0	0	0	0	0.1959	0
37777	Liver	1584	30687	19703	9184	59574	8367	7407	3044	18818	0	0	0	0	0.2401	0
	Muscle	800.4	47593	24071	6679	78343	14617	5658	1527	21802	0	0	0	0	0.2177	0
37778	Liver	1584	39996	26008	11795	77799	9730	6939	3671	20340	0	0	0	0	0.2073	0
	Muscle	800.4	103770	51011	14092	168873	29062	12595	3482	45139	0	0	0	0	0.2109	0
37929	Liver	1584	28247	16817	7991	53055	7207	4407	2556	14170	0	0	0	0	0.2108	0
	Muscle	800.4	89564	41624	12278	143466	24680	11459	3564	39703	0	0	0	0	0.2168	0
37930	Urine		273164	250899	133331	657394	329550	273162	145010	747722	6339	2049	468	8856	0.5321	0.0133
37931	Urine		29398	27546	13857	70801	2532	1039	536	4107	15372	12523	6220	34115	0.0548	0.3252
Unrelated6																
38243	Liver	1584	88664	67445	29411	185520	17190	11367	5487	34044	13140	8382	4813	26335	0.1551	0.1243
	Muscle	800.4	38222	17138	6049	61409	7305	2671	1266	11242	5354	2283	905	8542	0.1547	0.1221

38244	Liver	1584	62636	41373	19520	123529	11191	5713	2645	19549	13577	8671	4015	26263	0.1366	0.1753
	Muscle	800.4	25676	10926	4080	40682	2903	1229	311	4443	4314	1598	322	6234	0.0985	0.1329
38245	Liver	1584	77437	41404	15530	134371	11772	7162	2502	21436	14860	7855	3454	26169	0.1376	0.163
	Muscle	800.4	44271	20141	5232	69644	5617	1689	674	7980	5791	2124	627	8542	0.1028	0.1093
38246	Liver	1584	88667	57788	24326	170781	15148	9202	4241	28591	14584	9554	4661	28799	0.1434	0.1443
	Muscle	800.4	41610	16445	5455	63510	5381	3023	477	8881	5329	2327	708	8364	0.1227	0.1164
38247	Liver	1584	72193	48590	20941	141724	11698	7239	3157	22094	12792	8088	4026	24906	0.1349	0.1495
	Muscle	800.4	40577	17909	4327	62813	4822	2235	605	7662	6880	2875	1005	10760	0.1087	0.1462
38248	Liver	1584	61842	40168	16708	118718	10565	6447	3907	20919	12845	6783	3406	23034	0.1498	0.1625
	Muscle	800.4	56923	26613	7386	90922	9552	4160	1483	15195	12011	5019	1214	18244	0.1432	0.1671
38249	Liver	1584	45030	26625	11395	83050	10242	5768	2838	18848	7712	6102	2084	15898	0.185	0.1607
	Muscle	800.4	30307	13233	4129	47669	5671	2952	1013	9636	5696	2488	1252	9436	0.1682	0.1652
38250	Liver	1584	51242	24259	9711	85212	8782	4567	2118	15467	8816	5110	1963	15889	0.1536	0.1572
	Muscle	800.4	27867	13023	2967	43857	4867	2114	763	7744	4443	1822	1082	7347	0.1501	0.1435
38251	Liver	1584	16726	7808	3129	27663	3714	1930	823	6467	3774	1712	597	6083	0.1895	0.1803
	Muscle	800.4	31327	14011	3680	49018	6184	2493	822	9499	5639	2744	687	9070	0.1623	0.1561

38252	Liver	1584	98086	69202	32432	199720	16363	11145	5467	32975	18972	9992	6579	35543	0.1417	0.1511
	Muscle	800.4	42082	19717	5942	67741	5116	2529	818	8463	6443	3370	1056	10869	0.1111	0.1383
38253	Liver	1584	68791	42003	17555	128349	12352	6875	3347	22574	16308	8986	4725	30019	0.1496	0.1896
	Muscle	800.4	35046	17042	5047	57135	5304	2256	632	8192	5708	2392	996	9096	0.1254	0.1373
38254	Urine		24936	25701	15211	65848	3183	1823	661	5667	23117	23863	12901	59881	0.0792	0.4763
38255	Urine		76186	66400	32548	175134	62475	51415	26031	139921	1302	625	338	2265	0.4441	0.0128
NC3																
38256	Urine		55332	52119	29578	137029	11641	7411	5115	24167	32050	29150	14845	76045	0.1499	0.3569
38257	Urine		343051	288725	160198	791974	311852	259271	130096	701219	6574	3070	912	10556	0.4696	0.0132
38258	Liver	1584	17849	9631	4627	32107	4637	2665	1335	8637	0	0	0	0	0.212	0
	Muscle	800.4	82608	40683	11326	134617	24396	10723	3143	38262	0	0	0	0	0.2213	0
38259	Liver	1584	16322	10410	5427	32159	5292	2655	1384	9331	0	0	0	0	0.2249	0
	Muscle	800.4	129787	61055	18523	209365	33999	16303	4853	55155	0	0	0	0	0.2085	0
NC4																
38261	Liver	1584	21336	14452	7471	43259	8056	4563	2300	14919	0	0	0	0	0.2564	0
	Muscle	800.4	107286	54060	15890	177236	29301	13369	5495	48165	0	0	0	0	0.2137	0

38262	Liver	1584	19294	12855	5744	37893	5917	3845	2368	12130	0	0	0	0	0.2425	0
	Muscle	800.4	50829	23428	6571	80828	14903	5560	2093	22556	0	0	0	0	0.2182	0
38263	Liver	1584	20150	12467	5584	38201	6726	4369	1328	12423	0	0	0	0	0.2454	0
	Muscle	800.4	69767	35593	9786	115146	20086	8480	3052	31618	0	0	0	0	0.2154	0
38264	Urine		219500	209533	115007	544040	18283	6532	2802	27617	216404	196971	104015	517390	0.0483	0.4874
38265	Urine		68510	64718	33741	166969	78209	69724	35620	183553	7248	5117	2562	14927	0.5237	0

Supplementary B: Communal nursing study

Spreadsheet where the peptide intensities of the high turnover pup liver and muscle proteins and the mother's urine were recorded, along with the intensities of the corresponding labelled peptides, where mothered were fed either a [D_4] or [D_9] lysine labelled diet over the course of the experiment. Precursor RIA values were calculated from this data. Pup IDs/mother IDs in red are pups from the [D_9] lysine labelled mother/mothers fed [D_9] lysine labelled diet. Those in green are the pups from the [D_4] lysine labelled mother/mothers fed [D_4] lysine labelled diet.

ID:	Sample type	Peptide mass	Light peptide				D4 peptide				D9 peptide				IH/(IH+IL)	
Unrelated1			1	2	3	Sum	1	2	3	Sum	1	2	3	Sum	D4	D9
36286	Liver	916.5	69300	45070	17168	131538	12507	8652	3813	24972	17231	11735	4872	33838	0.1596	0.205
	Muscle	871.5	214736	122055	42586	379377	40983	20763	10164	71910	55122	29688	8389	93199	0.1593	0.197
36287	Liver	916.5	23401	16132	7436	46969	5420	3025	1596	10041	5557	3605	1844	11006	0.1761	0.19
	Muscle	871.5	68211	35027	10416	113654	12580	6984	2077	21641	14998	6955	2137	24090	0.16	0.175
36288	Liver	916.5	23298	16446	6164	45908	5128	3899	1484	10511	5923	3490	1632	11045	0.1863	0.194
	Muscle	871.5	177440	90784	26974	295198	28782	14255	4895	47932	30926	15498	4504	50928	0.1397	0.147
36289	Liver	916.5	46874	30975	12560	90409	10604	6040	2874	19518	11057	7377	3173	21607	0.1776	0.193
	Muscle	871.5	192285	93997	28459	314741	32687	16492	5504	54683	42136	21104	6043	69283	0.148	0.18
36290	Liver	916.5	23430	16218	5684	45332	4514	3242	1460	9216	6274	4018	1666	11958	0.169	0.209
	Muscle	871.5	164882	80955	26659	272496	23099	13312	3959	40370	30855	14121	4030	49006	0.129	0.152
36291	Liver	916.5	58442	40175	15719	114336	12343	7899	3726	23968	14186	9924	3853	27963	0.1733	0.197
	Muscle	871.5	269366	134045	39313	442724	42527	21699	8774	73000	54991	27063	9562	91616	0.1415	0.171
36292	Liver	916.5	22943	14547	6244	43734	5657	3615	1568	10840	5489	3560	1388	10437	0.1986	0.193
	Muscle	871.5	268079	136576	39952	444607	48841	28694	8857	86392	54848	27256	8717	90821	0.1627	0.17

36293	Liver	916.5	48744	33231	13164	95139	9056	5992	3009	18057	10221	6637	3164	20022	0.1595	0.174
	Muscle	871.5	152712	77663	23999	254374	25086	13638	4075	42799	34875	17378	4801	57054	0.144	0.183
36294	Liver	916.5	21649	15239	5316	42204	5133	2828	1463	9424	5247	3225	1363	9835	0.1825	0.189
	Muscle	871.5	53699	29294	9701	92694	10033	6062	1825	17920	12410	5738	1693	19841	0.162	0.176
36295	Liver	916.5	47337	32433	12795	92565	11195	7369	3918	22482	13556	8231	3752	25539	0.1954	0.216
	Muscle	871.5	92567	48242	14536	155345	16774	9209	4096	30079	16978	8197	2619	27794	0.1622	0.152
36296	Urine		36217	55644	31022	122883	67918	55667	31345	154930	2368	1144	334	3846	0.5577	0.03
36297	Urine		14458	13333	8168	35959	2090	665	0	2755	15286	16236	8733	40255	0.0712	0.528
Related 1																
36329	Liver	916.5	33722	22695	8959	65376	6023	3941	2154	12118	6932	4709	1753	13394	0.1564	0.17
	Muscle	871.5	989867	477360	142045	2E+06	1E+05	63754	26817	217607	160496	76887	21612	258995	0.1191	0.139
36330	Liver	916.5	18321	12557	4616	35494	3395	2607	1229	7231	4169	2789	1086	8044	0.1692	0.185
	Muscle	871.5	214418	106473	32163	353054	28407	14650	5236	48293	37240	16481	5699	59420	0.1203	0.144
36331	Liver	916.5	46999	30741	11950	89690	7345	5393	1915	14653	10093	6369	2596	19058	0.1404	0.175
	Muscle	871.5	1E+06	533501	146760	2E+06	1E+05	59877	19967	194344	169224	79215	23043	271482	0.0976	0.131

36332	Liver	916.5	20515	13359	5011	38885	3480	1789	866	6135	2929	2164	715	5808	0.1363	0.13
	Muscle	871.5	632812	31248	88643	1E+06	70328	37902	11489	119719	81419	38091	11505	131015	0.1038	0.112
36333	Liver	916.5	16698	10871	4163	31732	3452	2473	1252	7177	2942	2111	959	6012	0.1845	0.159
	Muscle	871.5	445048	21953	62226	726811	58393	32422	10314	101129	68565	32766	9709	111040	0.1221	0.133
36334	Liver	916.5	10543	7067	2975	20585	2088	1245	503	3836	1141	1298	790	3229	0.1571	0.136
	Muscle	871.5	262425	130762	39272	432459	36211	19641	6531	62383	39028	17765	5649	62442	0.1261	0.126
36335	Liver	916.5	15560	11319	3891	30770	3036	1862	979	5877	3068	1951	925	5944	0.1604	0.162
	Muscle	871.5	406609	195806	58719	661134	50065	26226	8889	85180	65968	31449	9511	106928	0.1141	0.139
36336	Liver	916.5	9112	6549	2513	18174	1816	772	488	3076	1388	1230	445	3063	0.1448	0.144
	Muscle	871.5	423659	197974	61733	683366	51326	27611	9166	88103	66324	33697	9407	109428	0.1142	0.138
36337	Liver	916.5	16016	11183	4425	31624	3239	2168	794	6201	2810	2291	900	6001	0.1639	0.159
	Muscle	871.5	612569	302087	85012	999668	77374	42710	12926	133010	76322	35252	9438	121012	0.1174	0.108
36338	Liver	916.5	8695	6085	2568	17348	1851	1111	688	3650	1697	1142	578	3417	0.1738	0.165
	Muscle	871.5	610754	387493	84151	1E+06	77360	40161	12364	129885	98655	48955	13703	161313	0.1071	0.13
36339	Urine		73649	65705	34548	173902	73755	63210	31006	167971	1888	1125	592	3605	0.4913	0.02
36340	Urine		80886	72295	36681	189862	5118	2832	775	8725	30326	25103	12477	67906	0.0439	0.263

Related 2																
36472	Liver	916.5	3274	1893	921	6088	3042	1354	833	5229	554	403	249	1206	0.462	0.165
36473	Liver	916.5	27957	18002	7440	53399	6623	4636	1881	13140	3326	2463	1013	6802	0.1975	0.113
	Muscle	871.5	232634	116616	35702	384952	44808	22256	8355	75419	23141	12464	4192	39797	0.1638	0.094
36474	Liver	916.5	23002	16725	6481	46208	7197	4188	1918	13303	3159	2506	891	6556	0.2235	0.124
	Muscle	871.5	114843	55923	17738	188504	30122	13866	4637	48625	12132	6454	1972	20558	0.2051	0.098
36475	Liver	916.5	23398	15245	6232	44875	6123	4309	1699	12131	2956	1865	919	5740	0.2128	0.113
	Muscle	871.5	323016	158871	48938	530825	71846	36476	10156	118478	33162	16680	5907	55749	0.1825	0.095
36476	Urine		22959	23281	11741	57981	27405	22923	11804	62132	2026	1652	770	4448	0.5173	0.071
36477	Urine		2329	1956	1115	5400	473	188	127	788	1208	1133	646	2987	0.1273	0.356
Unrelated2																
36679	Liver	916.5	10434	6711	2933	20078	2014	1232	490	3736	2106	998	449	3553	0.1569	0.15
	Muscle	871.5	46254	23153	6908	76315	8043	3769	994	12806	9415	5441	1524	16380	0.1437	0.177
36680	Liver	916.5	17404	11296	4838	33538	4241	2336	1448	8025	4612	3289	1290	9191	0.1931	0.215
	Muscle	871.5	1017	593	387	1997	281	111	66	458	419	158	19	596	0.1866	0.23

36681	Liver	916.5	13156	8421	3555	25132	2321	1804	715	4840	2665	1471	847	4983	0.1615	0.165
	Muscle	871.5	112422	55802	16409	184633	20740	10977	3737	35454	23905	14151	3734	41790	0.1611	0.185
36682	Liver	916.5	25004	16048	6914	47966	6624	3407	1637	11668	5004	3132	1400	9536	0.1957	0.166
	Muscle	871.5	52257	26601	8019	86877	10174	5737	1691	17602	11205	6501	1889	19595	0.1685	0.184
36683	Liver	916.5	16148	10581	4459	31188	4349	2872	1135	8356	4102	2928	860	7890	0.2113	0.202
	Muscle	871.5	125432	62455	16264	204151	24080	12049	3999	40128	24132	11795	3701	39628	0.1643	0.163
36684	Liver	916.5	10860	6801	3113	20774	2502	1630	792	4924	2581	1856	659	5096	0.1916	0.197
	Muscle	871.5	104066	51047	15667	170780	18410	9233	2854	30497	18567	10808	3304	32679	0.1515	0.161
36685	Liver	916.5	2969	2490	820	6279	797	505	203	1505	722	381	283	1386	0.1933	0.181
	Muscle	871.5	64028	30381	9688	104097	11858	5902	1563	19323	12339	5788	1873	20000	0.1566	0.161
36686	Liver	916.5	12095	7445	3309	22849	3043	1627	447	5117	2425	2130	840	5395	0.183	0.191
	Muscle	871.5	140625	69682	18882	229189	22519	11953	3515	37987	30870	15364	3959	50193	0.1422	0.18
36687	Liver	916.5	2091	1246	439	3776	522	540	167	1229	623	308	321	1252	0.2456	0.249
	Muscle	871.5	12731	6936	2085	21752	1908	1347	465	3720	2702	2083	508	5293	0.146	0.196
36478	Urine		83234	80404	46570	210208	13687	8080	4270	26037	84082	79480	45196	208758	0.1102	0.498
36688	Urine		17271	14608	8596	40475	20187	18652	10056	48895	764	298	177	1239	0.5471	0.03

Related 3																
37251	Liver	916.5	7342	4787	2131	14260	342	100	69	511	76	24	21	121	0.0346	0.008
	Muscle	871.5	22762	9969	3342	36073	2602	1245	400	4247	3374	1894	366	5634	0.1053	0.135
37252	Liver	916.5	16004	10633	4737	31374	3065	2587	1028	6680	2854	2486	987	6327	0.1755	0.168
	Muscle	871.5	106946	53706	18172	178824	16636	8928	3590	29154	20194	11397	3103	34694	0.1402	0.162
37253	Liver	916.5	19326	12145	5266	36737	3122	2091	917	6130	3805	2996	994	7795	0.143	0.175
	Muscle	871.5	227803	115506	35246	378555	28799	15206	5977	49982	41021	22116	5974	69111	0.1166	0.154
37254	Liver	916.5	42446	27708	11743	81897	6449	4137	1673	12259	8165	5564	2244	15973	0.1302	0.163
	Muscle	871.5	270000	134235	43478	447713	34230	18042	5596	57868	58383	30428	9289	98100	0.1145	0.18
37255	Liver	916.5	16312	11377	4508	32197	2836	1560	694	5090	2646	2150	745	5541	0.1365	0.147
	Muscle	871.5	331361	160635	46721	538717	37266	16554	5744	59564	57751	31174	7979	96904	0.0996	0.152
37256	Liver	916.5	24730	17518	6569	48817	3371	2413	1086	6870	5048	3366	983	9397	0.1234	0.161
	Muscle	871.5	92689	44440	13169	150298	9405	4825	1896	16126	16690	8073	2148	26911	0.0969	0.152
37257	Liver	916.5	20159	12510	4899	37568	3326	2184	1029	6539	3547	2089	1109	6745	0.1483	0.152
	Muscle	871.5	37453	18334	6331	62118	6345	3805	969	11119	6220	3319	998	10537	0.1518	0.145

37258	Urine		119132	107056	63157	289345	8576	2807	991	12374	109625	98541	54432	262598	0.041	0.476
37259	Urine		159738	144716	75066	379520	2E+05	1E+05	64363	339827	3628	2320	880	6828	0.4724	0.018
Related 4																
37260	Liver	916.5	55722	37006	15106	107834	11106	7079	3130	21315	13623	9409	3348	26380	0.165	0.197
	Muscle	871.5	129062	60392	19895	209349	20897	10753	3024	34674	28902	15200	4659	48761	0.1421	0.189
37261	Liver	916.5	54918	36041	16761	107720	9830	7415	3030	20275	17715	12432	4921	35068	0.1584	0.246
	Muscle	871.5	110271	58201	17140	185612	16443	8928	2877	28248	23184	12226	3324	38734	0.1321	0.173
37262	Liver	916.5	22333	14228	7432	43993	4645	3140	1463	9248	5519	3808	1308	10635	0.1737	0.195
	Muscle	871.5	121858	59042	16963	197863	25440	11771	3876	41087	25600	14165	3919	43684	0.1719	0.181
37263	Liver	916.5	18741	11761	5920	36422	2643	1515	912	5070	4350	3180	1125	8655	0.1222	0.192
	Muscle	871.5	217304	112885	32672	362861	23923	13009	4069	41001	51381	25111	6681	83173	0.1015	0.186
37264	Liver	916.5	12262	8701	3238	24201	2722	2207	832	5761	3060	2165	1390	6615	0.1923	0.215
	Muscle	871.5	99081	51000	15988	166069	18898	9381	3005	31284	28818	13234	4736	46788	0.1585	0.22
37265	Liver	916.5	14500	9609	4757	28866	2488	1840	796	5124	2980	2215	1163	6358	0.1508	0.181
	Muscle	871.5	128024	61615	18630	208269	16614	8516	2789	27919	25658	13670	3535	42863	0.1182	0.171

37266	Liver	916.5	23317	15596	7022	45935	5303	3408	2030	10741	5512	3930	1799	11241	0.1895	0.197
	Muscle	871.5	104510	53763	15132	173405	20275	10416	3780	34471	22741	12044	3322	38107	0.1658	0.18
37267	Liver	916.5	15354	10620	5801	31775	4400	2686	1128	8214	2975	2842	1177	6994	0.2054	0.18
	Muscle	871.5	77798	38510	13516	129824	14777	8138	1868	24783	14456	8376	1644	24476	0.1603	0.159
37268	Liver	916.5	27795	18587	7300	53682	5619	4470	1943	12032	7396	5175	1890	14461	0.1831	0.212
	Muscle	871.5	184985	96253	30687	311925	32784	16804	5513	55101	44772	21935	5938	72645	0.1501	0.189
37269	Liver	916.5	21502	14440	5533	41475	4456	2584	1158	8198	5807	3669	1825	11301	0.165	0.214
	Muscle	871.5	26686	13890	4010	44586	5907	3432	1320	10659	6973	3032	887	10892	0.1929	0.196
37270	Liver	916.5	120648	77948	31252	229848	22653	14536	6073	43262	27135	18689	7144	52968	0.1584	0.187
	Muscle	871.5	13383	6697	2429	22509	1722	1322	699	3743	2368	1367	1504	5239	0.1426	0.189
37271	Liver	916.5	211717	106048	32439	350204	31996	15475	5067	52538	43490	21789	5891	71170	0.1305	0.169
	Muscle	871.5	150884	74916	24462	250262	23548	13282	4632	41462	36454	17893	5458	59805	0.1421	0.193
37272	Liver	916.5	52119	37560	14171	103850	11492	8585	4629	24706	16818	10689	3879	31386	0.1922	0.232
	Muscle	871.5	251830	132446	39900	424176	34306	18444	6004	58754	51717	24850	7384	83951	0.1217	0.165
37273	Liver	916.5	37184	26502	12276	75962	11256	6424	3020	20700	11581	6636	2227	20444	0.2141	0.212
	Muscle	871.5	50176	26411	8546	85133	10296	5045	2358	17699	10600	5575	1280	17455	0.1721	0.17

37274	Liver	916.5	23291	13173	5335	41799	5659	3443	1437	10539	5158	3352	1251	9761	0.2014	0.189
	Muscle	871.5	163377	81467	26425	271269	31906	16556	5762	54224	30609	15168	4868	50645	0.1666	0.157
37275	Urine		64928	64806	36597	166331	5803	2111	692	8606	63795	64152	37653	165600	0.0492	0.499
37276	Urine		45610	41195	22179	108984	53178	46278	25050	124506	1827	911	545	3283	0.5332	0.029
Related 5																
37293	Liver	916.5	112988	72678	28866	214532	30537	18633	8164	57334	22328	14884	6695	43907	0.2109	0.17
	Muscle	871.5	870010	439023	132989	1E+06	2E+05	80982	23821	255928	133620	65071	18624	217315	0.1507	0.131
37294	Liver	916.5	146873	101591	39972	288436	39443	24683	12883	77009	26929	19475	7774	54178	0.2107	0.158
	Muscle	871.5	779409	387843	116030	1E+06	1E+05	71480	21559	233082	119510	56879	14740	191129	0.1537	0.13
37295	Liver	916.5	138216	93314	35982	267512	41706	26539	11511	79756	24324	15420	5699	45443	0.2297	0.145
	Muscle	871.5	479979	238768	71984	790731	1E+05	58844	18888	187316	80595	38516	11724	130835	0.1915	0.142
37296	Liver	916.5	101448	68135	26470	196053	24593	15479	6528	46600	21347	16172	6421	43940	0.192	0.183
	Muscle	871.5	446424	225485	68046	739955	71824	38632	12473	122929	68444	33917	10590	112951	0.1425	0.132
37297	Liver	916.5	31535	19852	8265	59652	8353	6634	2553	17540	7625	4833	1751	14209	0.2272	0.192
	Muscle	871.5	498894	257676	72686	829256	1E+05	60816	18339	197611	89883	43648	13076	146607	0.1924	0.15

37298	Liver	916.5	91633	62719	24346	178698	25770	17400	6555	49725	17557	11241	4246	33044	0.2177	0.156
	Muscle	871.5	444082	221101	63464	728647	1E+05	52820	17184	170384	70033	33659	10877	114569	0.1895	0.136
37299	Liver	916.5	43508	28552	11123	83183	10881	7195	3631	21707	9431	6252	2630	18313	0.207	0.18
	Muscle	871.5	266788	133418	40580	440786	54767	27994	10250	93011	55104	29390	8037	92531	0.1742	0.174
37300	Liver	916.5	46237	30106	13360	89703	16459	10554	4580	31593	10666	7106	3155	20927	0.2605	0.189
	Muscle	871.5	409655	208745	59848	678248	96485	52918	15808	165211	67306	33362	9960	110628	0.1959	0.14
37301	Liver	916.5	140310	91932	37242	269484	39837	26162	10162	76161	24108	18622	7410	50140	0.2203	0.157
	Muscle	871.5	882660	435249	126020	1E+06	1E+05	87605	24741	262001	114903	53172	14998	183073	0.1536	0.113
37302	Liver	916.5	39141	26095	11096	76332	10131	7825	3058	21014	9340	6327	3180	18847	0.2159	0.198
	Muscle	871.5	241712	117326	34367	393405	47257	27241	8073	82571	43782	21504	6182	71468	0.1735	0.154
37303	Liver	916.5	27896	18518	8833	55247	9491	6153	2698	18342	10335	5920	2313	18568	0.2492	0.252
	Muscle	871.5	446567	220449	65771	732787	98634	52175	17433	168242	81223	39995	12513	133731	0.1867	0.154
37304	Urine		26835	24393	12588	63816	29773	25686	14650	70109	771	359	0	1130	0.5235	0.017
37305	Urine		2164	2415	1227	5806	523	154	0	677	1250	1636	1183	4069	0.1044	0.412
NC1																
37306	Liver	916.5	9168	5516	2354	17038	0	0	0	0	1890	2478	499	4867	0	0.222

37307	Liver	916.5	14544	8948	3946	27438	0	0	0	0	3888	2212	809	6909	0	0.201
37308	Liver	916.5	13552	9185	3619	26356	0	0	0	0	3195	2411	846	6452	0	0.197
37309	Liver	916.5	24384	16310	6291	46985	0	0	0	0	6564	3682	1647	11893	0	0.202
37310	Liver	916.5	2847	1630	726	5203	0	0	0	0	627	462	156	1245	0	0.193
37311	Liver	916.5	16468	9814	4154	30436	0	0	0	0	3118	2465	1179	6762	0	0.182
37312	Liver	916.5	15981	9784	4250	30015	0	0	0	0	3483	2678	715	6876	0	0.186
37313	Urine		56432	55944	32232	144608	5271	1474	930	7675	63467	60383	34780	158630	0.0504	0.523
37314	Urine		216032	175443	79987	471462	49642	38531	18514	106687	8792	6445	3112	18349	0.1845	0.037
Related 6																
37316	Liver	916.5	12253	8530	3404	24187	2520	1950	506	4976	1949	1744	602	4295	0.1706	0.151
	Muscle	871.5	209269	105589	30464	345322	31227	17203	6212	54642	33694	15853	5347	54894	0.1366	0.137
37317	Liver	916.5	26969	19297	7601	53867	4845	3013	1421	9279	7686	5408	2171	15265	0.1469	0.221
	Muscle	871.5	690618	328572	92961	1E+06	70953	37149	11540	119642	135229	63831	19713	218773	0.0971	0.164
37318	Liver	916.5	28530	17849	7067	53446	5378	4129	1271	10778	6654	4294	1891	12839	0.1678	0.194
	Muscle	871.5	768739	372868	110286	1E+06	91571	46829	18253	156653	131049	64457	17651	213157	0.1112	0.145

37319	Liver	916.5	12787	7376	3207	23370	2781	1527	550	4858	2490	1807	635	4932	0.1721	0.174
	Muscle	871.5	300736	148456	46853	496045	44813	25631	8986	79430	63765	31797	10047	105609	0.138	0.176
37320	Liver	916.5	14450	9684	3327	27461	2313	1696	650	4659	4147	2730	764	7641	0.145	0.218
	Muscle	871.5	279839	134889	40795	455523	36745	19292	8959	64996	53837	30109	8929	92875	0.1249	0.169
37321	Liver	916.5	35611	23296	9250	68157	8902	5555	2397	16854	8056	5185	2131	15372	0.1983	0.184
	Muscle	871.5	243830	116908	37074	397812	45039	23697	10449	79185	45924	25842	7595	79361	0.166	0.166
37322	Liver	916.5	18386	11691	4792	34869	2780	1675	940	5395	4230	2798	906	7934	0.134	0.185
	Muscle	871.5	694509	321608	87812	1E+06	69310	36231	15448	120989	113473	62248	18094	193815	0.0988	0.149
37323	Liver	916.5	25646	18638	7036	51320	4300	3206	1423	8929	7144	4725	1802	13671	0.1482	0.21
	Muscle	871.5	186626	92258	27877	306761	26983	13500	6285	46768	36841	21707	7158	65706	0.1323	0.176
37324	Liver	916.5	12413	8345	3820	24578	1806	1544	781	4131	4023	3230	940	8193	0.1439	0.25
	Muscle	871.5	737349	345890	93495	1E+06	84349	40144	14074	138567	133290	72729	20370	226389	0.1054	0.161
37325	Liver	916.5	25572	17461	6060	49093	3335	2367	1232	6934	8300	5867	2202	16369	0.1238	0.25
	Muscle	871.5	276134	131336	38936	446406	28181	14177	4882	47240	62611	33448	10403	106462	0.0957	0.193
37326	Liver	916.5	4020	3006	1300	8326	1035	647	286	1968	1399	1032	437	2868	0.1912	0.256
	Muscle	871.5	901783	413611	121736	1E+06	1E+05	51385	19000	173777	170407	83124	23505	277036	0.1079	0.162

37327	Liver	916.5	16351	10990	4687	32028	3217	2409	1046	6672	2710	1612	794	5116	0.1724	0.138
	Muscle	871.5	154975	75438	22849	253262	23182	11663	3325	38170	22080	10912	3157	36149	0.131	0.125
37328	Liver	916.5	18789	12724	5683	37196	4268	2641	1052	7961	5295	3139	1353	9787	0.1763	0.208
	Muscle	871.5	360490	180157	57037	597684	57541	27253	12041	96835	81262	43583	11682	136527	0.1394	0.186
37729	Urine		157894	140435	75356	373685	24007	14881	6676	45564	63770	54604	29114	147488	0.1087	0.283
37730	Urine		36188	33251	18268	87707	36934	30802	14210	81946	981	381	218	1580	0.483	0.018
Unrelated3																
37731	Liver	916.5	113736	76068	29434	219238	28737	16599	6876	52212	21871	16098	6208	44177	0.1923	0.168
	Muscle	871.5	166168	77934	22630	266732	28802	14381	5741	48924	22586	12117	3464	38167	0.155	0.125
37732	Liver	916.5	31322	19583	7506	58411	7778	5259	2427	15464	5413	3374	1450	10237	0.2093	0.149
	Muscle	871.5	400850	193432	52049	646331	73572	38377	11283	123232	54925	29948	8746	93619	0.1601	0.127
37733	Liver	916.5	25720	16960	6285	48965	5467	2689	1445	9601	7176	4340	2072	13588	0.1639	0.217
	Muscle	871.5	39224	19066	6872	65162	6628	3726	1103	11457	10903	5195	1876	17974	0.1495	0.216
37734	Liver	916.5	36905	24578	9677	71160	7291	5148	1848	14287	7946	5982	2096	16024	0.1672	0.184
	Muscle	871.5	50579	26164	9131	85874	10393	4611	1749	16753	10552	5851	2059	18462	0.1632	0.177
37735	Liver	916.5	19914	14058	5751	39723	5023	4084	1281	10388	3769	2349	1063	7181	0.2073	0.153

	Muscle	871.5	320098	157386	49247	526731	59596	32282	10717	102595	48923	22133	6915	77971	0.163	0.129
37736	Liver	916.5	5924	5015	2274	13213	6598	3980	2013	12591	2094	1345	537	3976	0.4879	0.231
	Muscle	871.5	48865	22785	8861	80511	10125	4871	2529	17525	12188	5888	2558	20634	0.1788	0.204
37737	Liver	916.5	23724	15620	6308	45652	5509	3083	1508	10100	6755	3465	1576	11796	0.1812	0.205
	Muscle	871.5	33166	17320	6634	57120	8390	3931	1225	13546	10033	4532	1659	16224	0.1917	0.221
37738	Liver	916.5	22258	14534	5499	42291	4624	2361	1169	8154	4775	3505	1077	9357	0.1616	0.181
	Muscle	871.5	47243	23499	9374	80116	9345	4369	1916	15630	12552	6719	2254	21525	0.1632	0.212
37739	Liver	916.5	93595	59085	24820	177500	24391	16130	7226	47747	16494	11173	4562	32229	0.212	0.154
	Muscle	871.5	371809	182754	51926	606489	66302	37236	12131	115669	53830	25148	8740	87718	0.1602	0.126
37740	Liver	916.5	53006	34022	14601	101629	12085	7375	3515	22975	10585	6293	2267	19145	0.1844	0.159
	Muscle	871.5	270668	136597	37390	444655	42522	23629	7556	73707	41629	22730	5828	70187	0.1422	0.136
37741	Liver	916.5	40379	27103	10845	78327	7164	5195	1770	14129	10092	6471	2510	19073	0.1528	0.196
	Muscle	871.5	27382	12671	5108	45161	5047	2440	1226	8713	8156	3630	1208	12994	0.1617	0.223
37742	Liver	916.5	31463	20675	8725	60863	7714	4456	2354	14524	4210	3587	1237	9034	0.1927	0.129
	Muscle	871.5	784818	385784	110065	1E+06	1E+05	67846	22155	221769	106076	50658	15477	172211	0.1476	0.119
37743	Liver	916.5	31505	18830	7459	57794	6967	5049	1962	13978	4706	3056	1052	8814	0.1948	0.132

	Muscle	871.5	28493	14324	4988	47805	6095	3045	1057	10197	8155	4129	1337	13621	0.1758	0.222
37744	Liver	916.5	25584	17451	7811	50846	6345	3997	1383	11725	4695	3544	1143	9382	0.1874	0.156
	Muscle	871.5	54050	26307	10918	91275	9901	4961	2366	17228	15426	6648	2235	24309	0.1588	0.21
37745	Liver	916.5	31712	20450	7876	60038	5536	3036	1595	10167	9031	5687	2024	16742	0.1448	0.218
	Muscle	871.5	192732	93061	28571	314364	34573	19153	5319	59045	27452	14210	4913	46575	0.1581	0.129
37746	Urine		9742	8125	5066	22933	11163	9343	5162	25668	1052	265	0	1317	0.5281	0.054
37747	Urine		13778	13307	8340	35425	1500	569	224	2293	13887	11100	6896	31883	0.0608	0.474
Unrelated4																
37748	Liver	916.5	32424	22801	8449	63674	8382	6798	1970	17150	8392	5591	2340	16323	0.2122	0.204
	Muscle	871.5	33385	17203	5341	55929	6107	3467	1245	10819	7269	3561	1242	12072	0.1621	0.178
37749	Liver	916.5	21209	16296	6207	43712	7124	3828	1588	12540	5687	3563	1249	10499	0.2229	0.194
	Muscle	871.5	30548	15600	5070	51218	7548	3697	1157	12402	7664	3159	767	11590	0.1949	0.185
37750	Liver	916.5	31094	21264	8624	60982	8431	6163	2254	16848	7754	4278	1991	14023	0.2165	0.187
	Muscle	871.5	17894	8208	1927	28029	4029	2202	986	7217	2613	1344	692	4649	0.2048	0.142
37751	Liver	916.5	10010	6346	3159	19515	2812	1701	580	5093	1671	1602	464	3737	0.207	0.161
	Muscle	871.5	19460	10563	2759	32782	4910	2503	931	8344	4391	1594	408	6393	0.2029	0.163

37752	Liver	916.5	12760	9462	3362	25584	3936	2393	1021	7350	3208	2042	921	6171	0.2232	0.194
	Muscle	871.5	67631	37508	11445	116584	13584	7134	2153	22871	12903	6291	2127	21321	0.164	0.155
37622	Urine		72827	64542	33831	171200	70998	57720	29626	158344	2074	1728	387	4189	0.4805	0.024
37623	Urine		48927	43482	25069	117478	4840	1993	659	7492	52112	45654	27604	125370	0.06	0.516
Unrelated5																
37753	Liver	916.5	14944	10341	4034	29319	4757	3081	1601	9439	2573	2442	567	5582	0.2435	0.16
	Muscle	871.5	70668	36366	8859	115893	14699	8939	2271	25909	9265	4840	1556	15661	0.1827	0.119
37754	Liver	916.5	7503	4862	2361	14726	2426	1846	770	5042	2162	1429	868	4459	0.2551	0.232
	Muscle	871.5	3882	1436	678	5996	1439	810	341	2590	549	463	224	1236	0.3017	0.171
37755	Liver	916.5	8622	5256	2263	16141	2064	1657	556	4277	1390	1178	448	3016	0.2095	0.157
	Muscle	871.5	19418	9333	2788	31539	4935	3371	844	9150	3216	1782	661	5659	0.2249	0.152
37756	Liver	916.5	21735	14135	6807	42677	6997	5059	1875	13931	5062	3380	1420	9862	0.2461	0.188
	Muscle	871.5	13052	5807	2181	21040	2886	1760	764	5410	2388	1713	500	4601	0.2045	0.179
37757	Liver	916.5	19388	13501	5639	38528	6652	4650	1923	13225	3809	3126	1040	7975	0.2555	0.171
	Muscle	871.5	19295	10166	3422	32883	7006	3542	1394	11942	2896	1752	668	5316	0.2664	0.139
37758	Liver	916.5	7847	4393	2051	14291	3668	1967	616	6251	1337	1259	399	2995	0.3043	0.173

	Muscle	871.5	18532	10061	2599	31192	6057	3141	1123	10321	3561	2247	561	6369	0.2486	0.17
37759	Liver	916.5	28134	19979	7451	55564	9628	5389	3313	18330	5671	3827	1512	11010	0.2481	0.165
	Muscle	871.5	81737	39744	12615	134096	22443	13448	4195	40086	11297	6225	1450	18972	0.2301	0.124
37760	Liver	916.5	25967	20512	8604	55083	8208	5638	2041	15887	6376	3889	1841	12106	0.2239	0.18
	Muscle	871.5	46502	25238	6529	78269	14123	4808	1426	20357	7713	3890	1295	12898	0.2064	0.141
37761	Liver	916.5	10091	8010	2869	20970	3679	2409	817	6905	1772	1591	627	3990	0.2477	0.16
	Muscle	871.5	90765	45137	13993	149895	22165	11331	2984	36480	12093	5795	1473	19361	0.1957	0.114
37762	Liver	916.5	11855	8626	3154	23635	3388	2575	1213	7176	2144	1303	703	4150	0.2329	0.149
	Muscle	871.5	159330	78676	25449	263455	33292	16556	5232	55080	19106	9495	3167	31768	0.1729	0.108
37763	Urine		24860	19568	11599	56027	30288	23612	12953	66853	949	661	576	2186	0.5441	0.038
37764	Urine		17925	19312	10723	47960	2318	840	317	3475	18870	19602	10850	49322	0.0676	0.507
Related 7																
37765	Liver	916.5	16465	10302	4091	30858	4757	3698	932	9387	2424	1650	1118	5192	0.2332	0.144
	Muscle	871.5	40639	22303	6712	69654	9291	5078	1442	15811	5722	3805	943	10470	0.185	0.131
37766	Liver	916.5	11686	8397	2973	23056	2675	2752	1087	6514	1959	985	415	3359	0.2203	0.127

	Muscle	871.5	43191	21899	7158	72248	9413	5138	1281	15832	6064	2354	651	9069	0.1797	0.112
37767	Liver	916.5	29531	20047	8412	57990	8659	5065	2203	15927	4268	3074	1104	8446	0.2155	0.127
	Muscle	871.5	14740	7330	2607	24677	3136	1565	713	5414	2304	648	565	3517	0.1799	0.125
37768	Liver	916.5	18243	12344	4663	35250	5066	2883	1320	9269	2759	1709	608	5076	0.2082	0.126
	Muscle	871.5	46685	24124	6551	77360	9816	3584	1239	14639	7042	2962	1075	11079	0.1591	0.125
37769	Liver	916.5	25329	18486	6684	50499	6411	4937	1668	13016	3728	2131	1032	6891	0.2049	0.12
	Muscle	871.5	72519	35422	12114	120055	14384	8193	2696	25273	10069	5258	1264	16591	0.1739	0.121
37770	Liver	916.5	34874	22005	8560	65439	8703	5384	2080	16167	5748	3751	1566	11065	0.1981	0.145
	Muscle	871.5	168947	83819	24250	277016	33796	16290	5546	55632	17732	9890	3733	31355	0.1672	0.102
37771	Liver	916.5	12796	9180	3546	25522	3222	2123	1271	6616	2486	1374	1277	5137	0.2059	0.168
	Muscle	871.5	59907	31503	9767	101177	14596	6980	1922	23498	8665	3893	1611	14169	0.1885	0.123
37772	Liver	916.5	27349	19295	7233	53877	6446	4105	1585	12136	4706	3330	1339	9375	0.1838	0.148
	Muscle	871.5	62039	32783	10111	104933	10402	6638	2026	19066	8013	3611	931	12555	0.1538	0.107
37773	Liver	916.5	49662	30890	14202	94754	12286	7475	3294	23055	7037	4752	1675	13464	0.1957	0.124
	Muscle	871.5	68989	34291	9866	113146	16106	7465	2794	26365	7963	3899	1182	13044	0.189	0.103
37774	Urine		52888	48280	28207	129375	4124	1536	603	6263	44268	41670	25015	110953	0.0462	0.462

37775	Urine		91538	80196	43554	215288	1E+05	99352	51303	261652	2270	927	423	3620	0.5486	0.017
NC2																
37776	Liver	916.5	7057	4898	1614	13569	3009	2176	1146	6331	0	0	0	0	0.3181	0
	Muscle	871.5	915925	451635	130303	1E+06	3E+05	2E+05	46455	473132	0	0	0	0	0.24	0
37777	Liver	916.5	20777	14348	5313	40438	8598	6616	2838	18052	0	0	0	0	0.3086	0
	Muscle	871.5	233114	119192	38371	390677	91271	49987	17074	158332	0	0	0	0	0.2884	0
37778	Liver	916.5	29600	19549	10473	59622	14055	9723	3314	27092	0	0	0	0	0.3124	0
	Muscle	871.5	577704	288615	90182	956501	2E+05	1E+05	33222	336993	0	0	0	0	0.2605	0
37929	Liver	916.5	24151	15396	5395	44942	11042	8815	2701	22558	0	0	0	0	0.3342	0
	Muscle	871.5	657920	332268	103676	1E+06	2E+05	1E+05	38180	381967	0	0	0	0	0.2588	0
37930	Urine		247630	253028	134412	635070	3E+05	3E+05	1E+05	752998	6464	2063	468	8995	0.5425	0.014
37931	Urine		30290	28667	14614	73571	2743	1115	594	4452	15945	13037	6580	35562	0.0571	0.326
Unrelated6																
38243	Liver	916.5	53222	34107	13848	101177	15854	11076	4479	31409	10654	8730	3417	22801	0.2369	0.184
	Muscle	871.5	11471	6441	2011	19923	2245	1519	573	4337	2316	1049	993	4358	0.1788	0.179
38244	Liver	916.5	41249	27940	10423	79612	9826	6769	2714	19309	14250	9772	3759	27781	0.1952	0.259

	Muscle	871.5	8846	4756	1881	15483	1435	848	450	2733	2426	1036	1072	4534	0.15	0.227
38245	Liver	916.5	60330	40125	15777	116232	14243	10396	4147	28786	16645	12943	4866	34454	0.1985	0.229
	Muscle	871.5	21760	9705	3163	34628	3427	1860	560	5847	4715	1989	1038	7742	0.1445	0.183
38246	Liver	916.5	19406	14431	6021	39858	6510	3935	1765	12210	8301	5431	1926	15658	0.2345	0.282
	Muscle	871.5	17625	8531	3034	29190	3308	1936	466	5710	2613	1691	815	5119	0.1636	0.149
38247	Liver	916.5	32346	19303	8684	60333	8871	4228	1999	15098	8900	5724	2917	17541	0.2002	0.225
	Muscle	871.5	13404	6641	2434	22479	1929	826	509	3264	2638	1044	697	4379	0.1268	0.163
38248	Liver	916.5	30689	23106	8431	62226	9035	5778	3183	17996	11730	7164	2487	21381	0.2243	0.256
	Muscle	871.5	30226	15398	4668	50292	6180	2842	1197	10219	5468	2739	1995	10202	0.1689	0.169
38249	Liver	916.5	25180	16543	5898	47621	7372	5059	2442	14873	6817	4865	2165	13847	0.238	0.225
	Muscle	871.5	13186	7470	2442	23098	2448	1883	831	5162	2416	1260	1450	5126	0.1827	0.182
38250	Liver	916.5	29787	20955	7831	58573	6841	4821	2640	14302	11097	6963	3190	21250	0.1963	0.266
	Muscle	871.5	14475	7689	3261	25425	2293	1723	523	4539	3156	1377	838	5371	0.1515	0.174
38251	Liver	916.5	28940	19545	8191	56676	5639	4129	2787	12555	7295	5048	2220	14563	0.1813	0.204
	Muscle	871.5	11867	5787	1995	19649	2223	1667	469	4359	2189	941	1464	4594	0.1816	0.189
38252	Liver	916.5	53877	36077	12708	102662	13002	9370	3738	26110	18447	12315	4629	35391	0.2028	0.256

	Muscle	871.5	20250	11069	2701	34020	2785	1696	641	5122	3948	2317	1736	8001	0.1309	0.19
38253	Liver	916.5	48581	33525	12716	94822	12103	7891	4355	24349	16551	13344	5408	35303	0.2043	0.271
	Muscle	871.5	10592	6528	1804	18924	1757	1029	548	3334	2386	1417	1583	5386	0.1498	0.222
38254	Urine		25159	25945	15414	66518	3254	1921	747	5922	23215	24065	12934	60214	0.0818	0.475
38255	Urine		77155	67016	32880	177051	62955	51838	26373	141166	1363	336	356	2055	0.4436	0.011
NC3																
38256	Urine		55947	53003	29969	138919	7475	5484	1690	14649	32395	29564	15032	76991	0.0954	0.357
38257	Urine		364741	292004	161977	818722	3E+05	3E+05	1E+05	709033	6610	3220	969	10799	0.4641	0.013
38258	Liver	916.5	16987	11300	4089	32376	7197	4829	3183	15209	0	0	0	0	0.3196	0
	Muscle	871.5	755242	379696	111474	1E+06	3E+05	1E+05	44732	483412	0	0	0	0	0.2795	0
38259	Liver	916.5	15715	9900	3581	29196	6042	5678	1649	13369	0	0	0	0	0.3141	0
	Muscle	871.5	383568	186988	55842	626398	1E+05	71167	24361	238477	0	0	0	0	0.2757	0
NC4																
38261	Liver	916.5	19526	12081	2018	33625	9388	6258	2366	18012	0	0	0	0	0.3488	0
	Muscle	871.5	134942	72030	23929	230901	58930	28925	9973	97828	0	0	0	0	0.2976	0
38262	Liver	916.5	16817	10404	4672	31893	6613	5055	1804	13472	0	0	0	0	0.297	0

	Muscle	871.5	133824	64630	21136	219590	53658	25072	7789	86519	0	0	0	0	0.2826	0
38263	Liver	916.5	26344	17915	6981	51240	11094	8134	3746	22974	0	0	0	0	0.3096	0
	Muscle	871.5	140712	74867	25045	240624	57921	29933	9261	97115	0	0	0	0	0.2875	0
38264	Urine	916.5	220120	210368	115576	546064	18526	6666	2967	28159	216790	197760	104383	518933	0.049	0.487
38265	Urine	871.5	69080	65626	34301	169007	78955	70407	36226	18558	7283	5152	2605	15040	0.5234	0.082

Supplementary material B: Communal nursing study

Spreadsheet where the peptide intensities of the related mother's urine were recorded, along with the intensities of the corresponding labelled peptides, where mothered were fed either a [D_4] or [D_9] lysine labelled diet over the course of the experiment. Precursor RIA values were calculated from this data. Mother IDs in red are fed the [D_9] lysine labelled diet. Those in green fed the [D_4] lysine labelled diet.

Sample no.	Peptide	Light peptide				D4 peptide				D9 peptide				IH/(IH+IL)	
		1	2	3	Sum	1	2	3	Sum	1	2	3	Sum	D4	D9
36339	1569	73649	65705	34548	173902	73755	63210	31006	167971	1888	1125	592	3605	0.491	0.02
	2009	7539	8779	5663	21981	7235	7444	5272	19951	916	425	337	1678	0.476	0.071
	2741	1608	2831	1962	6401	1490	2383	1892	5765	216	124	114	454	0.474	0.066
36340	1569	80886	72295	36681	189862	5118	2832	775	8725	30326	25103	12477	67906	0.044	0.263
	2009	9291	12516	7029	28836	2070	160	0	2230	3654	4653	2795	11102	0.072	0.278
	2471	13763	20607	16843	51213	1378	514	214	2106	5321	6620	4619	16560	0.039	0.244
36476	1569	22959	23281	11741	57981	27405	22923	11804	62132	2026	1652	770	4448	0.517	0.071
	2009	2148	2357	1408	5913	1560	2067	1627	5254	337	292	123	752	0.47	0.113
	2471	982	1651	1282	3915	1312	1192	1261	3765	271	239	0	510	0.49	0.115
36477	1569	2329	1956	1115	5400	473	188	127	788	1208	1133	646	2987	0.127	0.356
	2009	413	556	368	1337	159	87	0	246	466	428	245	1139	0.155	0.46
37258	1569	119132	107056	63157	289345	8576	2807	991	12374	1E+05	98541	54432	262598	0.041	0.476
	2009	70757	88570	60938	220265	10724	4365	2688	17777	63947	78818	50864	193629	0.075	0.468
	2471	57822	90771	75067	223660	12161	4364	1178	17703	54478	78827	68243	201548	0.073	0.474
37259	1569	159738	144716	75066	379520	2E+05	1E+05	64363	339827	3628	2320	880	6828	0.472	0.018
	2009	4488	4779	4179	13446	4809	4707	3650	13166	231	195	0	426	0.495	0.031
	2471	7316	11061	9598	27975	8583	10638	6808	26029	962	576	278	1816	0.482	0.061
37275	1569	64928	64806	36597	166331	5803	2111	692	8606	63795	64152	37653	165600	0.049	0.499
	2009	3939	4534	3500	11973	722	194	44	960	3822	5207	3487	12516	0.074	0.511
	2471	9095	14511	14908	38514	2068	790	111	2969	11048	14457	13115	38620	0.072	0.501
37276	1569	45610	41195	22179	108984	53178	46278	25050	124506	1827	911	545	3283	0.533	0.029
	2009	247	333	304	884	224	307	168	699	0	0	0	0	0.442	0
	2471	255	556	489	1300	633	529	328	1490	0	0	0	0	0.534	0
37304	1569	26835	24393	12588	63816	29773	25686	14650	70109	771	359	0	1130	0.523	0.017
	2009	1740	1988	1343	5071	2162	2239	1541	5942	299	83	0	382	0.54	0.07

	2471	1385	1544	1784	4713	1772	2407	1426	5605	309	176	0	485	0.543	0.093
37305	1569	2164	2415	1227	5806	523	154	0	677	1250	1636	1183	4069	0.104	0.412
37729	1569	157894	140435	75356	373685	24007	14881	6676	45564	63770	54604	29114	147488	0.109	0.283
	2009	250	758	892	1900	133	75	21	229	228	280	220	728	0.108	0.277
	2471	398	673	310	1381	93	47	18	158	181	239	168	588	0.103	0.299
37730	1569	36188	33251	18268	87707	36934	30802	14210	81946	981	381	218	1580	0.483	0.018
	2009	427	852	565	1844	405	644	606	1655	86	0	0	86	0.473	0.045
	2471	1078	1521	1285	3884	1431	1278	1151	3860	221	82	0	303	0.498	0.072
37774	1569	52888	48280	28207	129375	4124	1536	603	6263	44268	41670	25015	110953	0.046	0.462
	2471	87	131	149	367	0	0	0	0	136	163	78	377	0	0.507
37775	1569	91538	80196	43554	215288	1E+05	99352	51303	261652	2270	927	423	3620	0.549	0.017
	2009	455	642	457	1554	510	829	553	1892	0	0	0	0	0.549	0
	2471	1329	1418	1586	4333	1879	1974	1836	5689	79	0	0	79	0.568	0.018

Supplementary material B: Communal nursing study

Spreadsheet where the peptide intensities of the unrelated mother's urine were recorded, along with the intensities of the corresponding labelled peptides, where mothered were fed either a [D_4] or [D_9] lysine labelled diet over the course of the experiment. Precursor RIA values were calculated from this data. Mother IDs in red are fed the [D_9] lysine labelled diet. Those in green fed the [D_4] lysine labelled diet.

Sample no.	Peptide	Light peptide				D4 peptide				D9 peptide				IH/(IH+IL)	
		1	2	3	Sum	1	2	3	Sum	1	2	3	Sum	D4	D9
36296	1569	36217	55644	31022	122883	67918	55667	31345	2E+05	2368	1144	334	3846	0.5577	0.0303
	2009	21654	27466	18151	67271	24046	27895	17835	69776	1360	449	394	2203	0.5091	0.0317
	2471	2327	2371	1564	6262	1941	2746	1735	6422	203	139	0	342	0.5063	0.0518
36297	1569	14458	13333	8168	35959	2090	665	0	2755	15286	16236	8733	40255	0.0712	0.5282
	2009	6206	8060	5566	19832	1222	156	99	1477	5226	6154	3828	15208	0.0693	0.434
	2471	1204	2043	1956	5203	282	130	84	496	1446	1784	1947	5177	0.087	0.4987
36478	1569	83234	80404	46570	210208	13687	8080	4270	26037	84082	79480	45196	208758	0.1102	0.4983
	2009	19350	23084	17029	59463	5106	3224	965	9295	18909	22250	15777	56936	0.1352	0.4891
	2471	27556	39615	33737	100908	8122	4215	2605	14942	26450	41095	33183	100728	0.129	0.4996
36688	1569	17271	14608	8596	40475	20187	18652	10056	48895	764	298	177	1239	0.5471	0.0297
	2009	1179	722	770	2671	883	744	659	2286	0	0	0	0	0.4612	0
	2471	442	542	294	1278	511	492	344	1347	0	0	0	0	0.5131	0
37746	1569	9742	8125	5066	22933	11163	9343	5162	25668	1052	265	0	1317	0.5281	0.0543
	1607	5516	6676	4269	16461	6700	6413	3848	16961	664	365	92	1121	0.5075	0.0638
					0				0				0	#DIV/0!	######
37747	1569	13778	13307	8340	35425	1500	569	224	2293	13887	11100	6896	31883	0.0608	0.4737
37622	1569	72827	64542	33831	171200	70998	57720	29626	2E+05	2074	1728	387	4189	0.4805	0.0239
	2009	31517	37274	24679	93470	28293	29617	18590	76500	4413	5096	2830	12339	0.4501	0.1166
	2471	17456	27557	22503	67516	24519	27356	18432	70307	3296	1554	965	5815	0.5101	0.0793
37623	1569	48927	43482	25069	117478	4840	1993	659	7492	52112	45654	27604	125370	0.06	0.5162
	2009	11006	13501	8307	32814	1274	703	621	2598	12108	14201	9186	35495	0.0734	0.5196
	2471	10577	18408	14391	43376	2478	873	866	4217	12236	20498	16047	48781	0.0886	0.5293
37763	1569	24860	19568	11599	56027	30288	23612	12953	66853	949	661	576	2186	0.5441	0.0376
	1607	13098	12436	8336	33870	14886	13196	7375	35457	1366	675	44	2085	0.5114	0.058

	2471	4385	4265	3339	11989	5595	5321	2613	13529	457	281	0	738	0.5302	0.058
37764	1569	17925	19312	10723	47960	2318	840	317	3475	18870	19602	10850	49322	0.0676	0.507
	2009	127	200	206	533	0	0	0	0	219	220	79	518	0	0.4929
	2471	134	188	251	573	0	0	0	0	213	174	192	579	0	0.5026
38254	1569	25159	25945	15414	66518	3254	1921	747	5922	23215	24065	12934	60214	0.0818	0.4751
	2009	2314	2380	1561	6255	297	420	145	862	1660	1872	1345	4877	0.1211	0.4381
	2471	694	982	467	2143	112	79	66	257	865	826	86	1777	0.1071	0.4533
38255	1569	77155	67016	32880	177051	62955	51838	26373	1E+05	1363	336	356	2055	0.4436	0.0115
	2009	12714	14108	10557	37379	10266	13332	8350	31948	1201	1151	614	2966	0.4608	0.0735
	2471	11231	14409	11328	36968	11380	13593	9204	34177	1114	586	0	1700	0.4804	0.044

Supplementary material B: Communal nursing study

Spreadsheet where the peptide intensities of the mother's urine were recorded (mothers with non-communal litters), along with the intensities of the corresponding labelled peptides, where mothered were fed either a [D_4] or [D_9] lysine labelled diet over the course of the experiment. Precursor RIA values were calculated from this data. Mother IDs in red are fed the [D_9] lysine labelled diet. Those in green fed the [D_4] lysine labelled diet.

Sample no.	Peptide	Light peptide				D4 peptide				D9 peptide				IH/(IH+IL)	
		1	2	3	Sum	1	2	3	Sum	1	2	3	Sum	D4	D9
37313	1569	56432	55944	32232	144608	5271	1474	930	7675	63467	60383	34780	158630	0.0504	0.5231
	2009	1910	2636	1475	6021	477	0	0	477	2470	2420	1571	6461	0.0734	0.5176
	2471	623	1089	996	2708	143	0	0	143	980	1002	707	2689	0.0502	0.4982
37314	1569	216032	175443	79987	471462	49642	38531	18514	106687	8792	6445	3112	18349	0.1845	0.0375
	2009	3993	3908	3028	10929	1211	1344	949	3504	349	249	218	816	0.2428	0.0695
	2471	5209	5734	4833	15776	1818	1889	1145	4852	296	391	202	889	0.2352	0.0533
37930	1569	247630	253028	134412	635070	331687	274916	1E+05	752998	6464	2063	468	8995	0.5425	0.014
	2009	7034	10056	6807	23897	9526	10496	8355	28377	722	464	210	1396	0.5429	0.0552
	2471	10500	12953	11593	35046	14792	17143	12563	44498	1659	490	322	2471	0.5594	0.0659
37931	1569	30290	28667	14614	73571	2743	1115	594	4452	15945	13037	6580	35562	0.0571	0.3259
	2009	421	357	569	1347	0	0	0	0	113	180	115	408	0	0.2325
	2471	263	194	450	907	0	0	0	0	194	205	86	485	0	0.3484
38256	1569	55947	53003	29969	138919	7475	5484	1690	14649	32395	29564	15032	76991	0.0954	0.3566
	2471	4528	7505	5661	17694	1555	732	391	2678	2986	4420	3753	11159	0.1315	0.3868

38257	1569	364741	292004	161977	818722	314887	262270	1E+05	709033	6610	3220	969	10799	0.4641	0.013
	2009	261157	306011	203246	770414	253981	276911	2E+05	710689	10475	4533	2129	17137	0.4798	0.0218
	2471	260459	393854	310029	964342	317474	374032	3E+05	965949	30589	13200	4569	48358	0.5004	0.0478
38264	1569	220120	210368	115576	546064	18526	6666	2967	28159	216790	197760	104383	518933	0.049	0.4873
	2009	168095	217782	148693	534570	29864	12574	0	42438	166996	200759	136622	504377	0.0735	0.4855
	2471	125581	210409	174796	510786	24692	9938	5715	40345	138888	181064	156206	476158	0.0732	0.4825
38265	1569	69080	65626	34301	169007	78955	70407	36226	185588	7283	5152	2605	15040	0.5234	0.0817
	2009	31539	36612	24298	92449	33426	38469	24415	96310	3501	3504	1704	8709	0.5102	0.0861
	2471	36034	51807	46591	134432	51762	62456	45591	159809	8394	5899	4592	18885	0.5431	0.1232

Supplementary material B: Communal nursing study – statistical analysis

- Tested for statistical significance between the mothers' investment in their own pups and the other pups (so, for example, compared the amount of investment a D_4 mother made in her own pups and the amount made in D_9's pups, and vice versa).
- Did this in R using analysis of variance (anova), which analyses the difference between group means along with variation among and between groups. If sample sizes were too small, Welch two sample t-test was used.
- The P value (Pr(>F)) is calculated from the anova table (f value, sum sq etc.). A low (>0.05) p value suggests that the difference between the two groups is down to more than just random sampling, and that the means are significantly different. These are highlighted in the tables below.
- Statistical significance between mothers' investment in the entire communal litter was tested for using Welch two sample t-test.

Related litters

Litter 1

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -0.1549, df = 77.351, p-value = 0.8773

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-0.01403034 0.01200492

sample estimates:

mean of x mean of y

0.1172974 0.1183101

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000424	0.0000424	0.105	0.8
D9 mother	1	8.5E-05	8.5E-05	0.274	0.693

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	1.49E-04	1.49E-04	2.605	0.353
D9 mother	1	1.8E-06	1.8E-06	0.632	0.572

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0007272	0.0007272	10.29	0.192
D9 mother	1	5.27E-05	5.27E-05	1.851	0.404

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)

D4 mother	1	0.0001175	0.0001175	0.766	0.542
D9 mother	1	1.19E-05	1.19E-05	0.521	0.602

Litter 2

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 4.3556, df = 16.783, p-value = 0.0004424

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 0.04796167 0.13825076

sample estimates:

mean of x mean of y

0.1936421 0.1005359

High turnover in liver

	t	df	p-value
D4 mother	11.9955	1.001	0.05277

D9 mother	1.6362	1.059	0.3389

Low turnover in liver

	t	df	p-value
D4 mother	5.5812	1.003	0.1124
D9 mother	5.3576	1.204	0.08743

Low turnover in muscle

	t	df	p-value
D4 mother	11.5353	1.979	0.0077
D9 mother	6.0117	1.01	0.1033

Litter 3

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -2.7058, df = 52.363, p-value = 0.009176

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-0.040163443 -0.005961826

sample estimates:

mean of x mean of y

0.1075265 0.1305892

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000247	0.0000247	0.0203	0.904
D9 mother	1	0.005891	0.005891	23.06	0.131

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.000314	0.000314	0.084	0.821
D9 mother	1	3.3E-05	3.3E-05	7.796	0.219

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)

D4 mother	1	6E-07	6E-07	0.012	0.93
D9 mother	1	0.0009295	0.0009295	9.224	0.202

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	3.32E-05	3.32E-05	0.581	0.585
D9 mother	1	0.00015	0.00015	1.143	0.479

Litter 4

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -3.5129, df = 117.3, p-value = 0.0006304

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -0.034294199 -0.009567269

sample estimates:

mean of x mean of y

0.1390210 0.1609517

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.00002	0.00002	0.02	0.895
D9 mother	1	3.5E-06	3.5E-06	0.009	0.93

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.000203	0.000203	0.383	0.57
D9 mother	1	0.0000014	0.0000014	0.008	0.932

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	7.91E0-6	7.91E0-6	0.226	0.66
D9 mother	1	0.0000077	0.0000077	0.017	0.901

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.000345	0.000345	3.185	0.149
D9 mother	1	1E-06	1E-06	0.008	0.933

Litter 5

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 4.4007, df = 85.341, p-value = 3.103e-05

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

0.01943695 0.05147200

sample estimates:

mean of x mean of y

0.1686868 0.1332323

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0003524	0.0003524	10.16	0.0859 ‘.’
D9 mother	1	2.29E-05	2.29E-05	0.154	0.733

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	2.45E-05	2.45E-05	0.825	0.46
D9 mother	1	0.0001681	0.0001681	6.591	0.124

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	3.09E-05	3.09E-05	0.789	0.468
D9 mother	1	0.0000105	0.0000105	0.02	0.901

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)

D4 mother	1	1.99E-05	1.99E-05	0.224	0.683
D9 mother	1	1.48E-05	1.48E-05	0.131	0.752

Litter 6

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -5.2086, df = 99.43, p-value = 1.03e-06

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -0.04975739 -0.02230628

sample estimates:

mean of x mean of y

0.1197495 0.1557814

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0005925	0.0005925	1.572	0.278
D9 mother	1	6.47E-05	6.47E-05	0.088	0.781

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000243	0.0000243	0.096	0.773
D9 mother	1	7.55E-05	7.55E-05	2.391	0.197

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0004386	0.0004386	4.593	0.0987 ‘.’
D9 mother	1	2.11E-06	2.11E-06	0.032	0.867

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	6.33E-04	6.33E-04	42.85	0.00282 ‘**’
D9 mother	1	9.07E-05	9.07E-05	2.203	0.212

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Litter 7

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 7.6728, df = 67.717, p-value = 8.81e-11

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 0.04345687 0.07400832

sample estimates:

mean of x mean of y

0.1641191 0.1053865

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000586	0.0000586	0.186	0.708
D9 mother	1	0.0004262	0.0004262	3.234	0.214

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0002012	0.0002012	6.904	0.119
D9 mother	1	0.0000038	0.0000038	0.017	0.908

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0004973	0.0004973	4.93	0.157
D9 mother	1	1.85E-04	1.85E-04	3.721	0.194

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000023	0.0000023	0.01	0.931
D9 mother	1	0.00065	0.00065	3.29	0.211

Unrelated litters

Litter 1

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -1.742, df = 74.723, p-value = 0.08562

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -0.024563261 0.001645902

sample estimates:

mean of x mean of y

0.1437995 0.1552582

High turnover in liver

	t	df			P value
D4 mother	-2.55	7.95			0.350
D9 mother	-1.07	4.58			0.338

High turnover in muscle

	t	df			P value
D4 mother	-5.08	6.55			0.175
D9 mother	-0.36	1.09			0.776

Low turnover in liver

	t	df			P value
D4 mother	-0.11	1.92			0.393
D9 mother	-0.21	3			0.850

Low turnover in muscle

	t	df			P value
D4 mother	-1.4	1.9			0.298
D9 mother	-1.01	1.09			0.487

Litter 2

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -1.3797, df = 69.352, p-value = 0.1721

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-0.027375339 0.004989537

sample estimates:

mean of x mean of y

0.1495456 0.1607385

High turnover in liver

	t	df			P value
D4 mother	1.6978	2.041			0.2292
D9 mother	0.1441	1.411			0.9032

High turnover in muscle

	t	df			P value
D4 mother	-1.8057	1.545			0.2483
D9 mother	-1.3518	1.116			0.3879

Low turnover in liver

	t	df			P value
D4 mother	0.9804	6.112			0.364

D9 mother	-0.5999	1.385			0.6318

Low turnover in muscle

	t	df			P value
D4 mother	-2.3223	6.62			0.05779
D9 mother	-1.4261	1.521			0.3247

Litter 3

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 0.4614, df = 113.197, p-value = 0.6454

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -0.01314853 0.02113337

sample estimates:

mean of x mean of y

0.1549398 0.1509474

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	7E-07	7E-07	0.005	0.945
D9 mother	1	0.03055	0.03055	2.659	0.164

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000134	0.0000134	0.057	0.82
D9 mother	1	0.0003539	0.0003539	4.862	0.079 ‘.’

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	7.19E-05	7.19E-05	0.927	0.38
D9 mother	1	0.000170	0.000170	1.092	0.344

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	3.80E-05	3.80E-05	0.289	0.614
D9 mother	1	0.000762	0.000762	5.465	0.066 ‘.’

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Litter 4

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 3.2676, df = 37.998, p-value = 0.002305

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 0.01139567 0.04850750

sample estimates:

mean of x mean of y

0.1822597 0.1523081

High turnover in liver

	t	df	P value
D4 mother	2.1904	2.637	0.1285
D9 mother	0.4854	2.114	0.6731

High turnover in muscle

	t	df	P value
D4 mother	0.5846	2.881	0.6014
D9 mother	0.986	1.722	0.4421

Low turnover in liver

	t	df	P value
D4 mother	2.1445	2.091	0.1596
D9 mother	0.7088	2.666	0.5354

Low turnover in muscle

	t	df	P value
D4 mother	1.192	2.018	0.3546
D9 mother	0.8715	2.632	0.4556

Litter 5

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 7.7648, df = 75.227, p-value = 3.285e-11

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

0.05192189 0.08775527

sample estimates:

mean of x mean of y

0.2035713 0.1337328

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0	0	0	0.994
D9 mother	1	0.000307	0.000307	0.345	0.617

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0002827	0.0002827	0.515	0.547
D9 mother	1	0.000648	0.000648	0.305	0.636

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.000712	0.000712	0.546	0.537
D9 mother	1	0.00072	0.00072	2.088	0.285

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0003876	0.0003876	9.171	0.0939 ‘.’
D9 mother	1	0.0004247	0.0004247	0.814	0.462

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Litter 6

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -2.1432, df = 81.871, p-value = 0.03507

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -0.035564382 -0.001323534

sample estimates:

mean of x mean of y

0.1628515 0.1812954

High turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000001	0.0000001	0	0.99
D9 mother	1	0.0003306	0.0003306	1.517	0.306

High turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0001897	0.0001897	0.28	0.633
D9 mother	1	0.0000975	0.0000975	0.208	0.68

Low turnover in liver

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0000195	0.0000195	0.036	0.862
D9 mother	1	0.0004966	0.0004966	1.209	0.352

Low turnover in muscle

	Df	Sum sq	Mean Sq	F value	Pr(>F)
D4 mother	1	0.0009762	0.0009762	1.258	0.344
D9 mother	1	0.001541	0.001541	3.326	0.166

Non Communal Litters

Litter 1

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = -18.4109, df = 13, p-value = 1.077e-10

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -0.1861130 -0.1470223

sample estimates:

mean of x mean of y

0.0000000 0.1665676

Litter 2

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 21.8468, df = 15, p-value = 8.743e-13

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:
0.2276477 0.2768703
sample estimates:
mean of x mean of y
0.252259 0.000000

Litter 3

MOTHERS INVESTMENT IN ENTIRE LITTER
Welch Two Sample t-test
data: data$D4 and data$D9
t = 15.8506, df = 7, p-value = 9.645e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
0.2186184 0.2952834
sample estimates:
mean of x mean of y
0.2569509 0.0000000

Litter 4

MOTHERS INVESTMENT IN ENTIRE LITTER

Welch Two Sample t-test

data: data$D4 and data$D9

t = 21.7305, df = 11, p-value = 2.19e-10

alternative hypothesis: true difference in means is not equal to 0

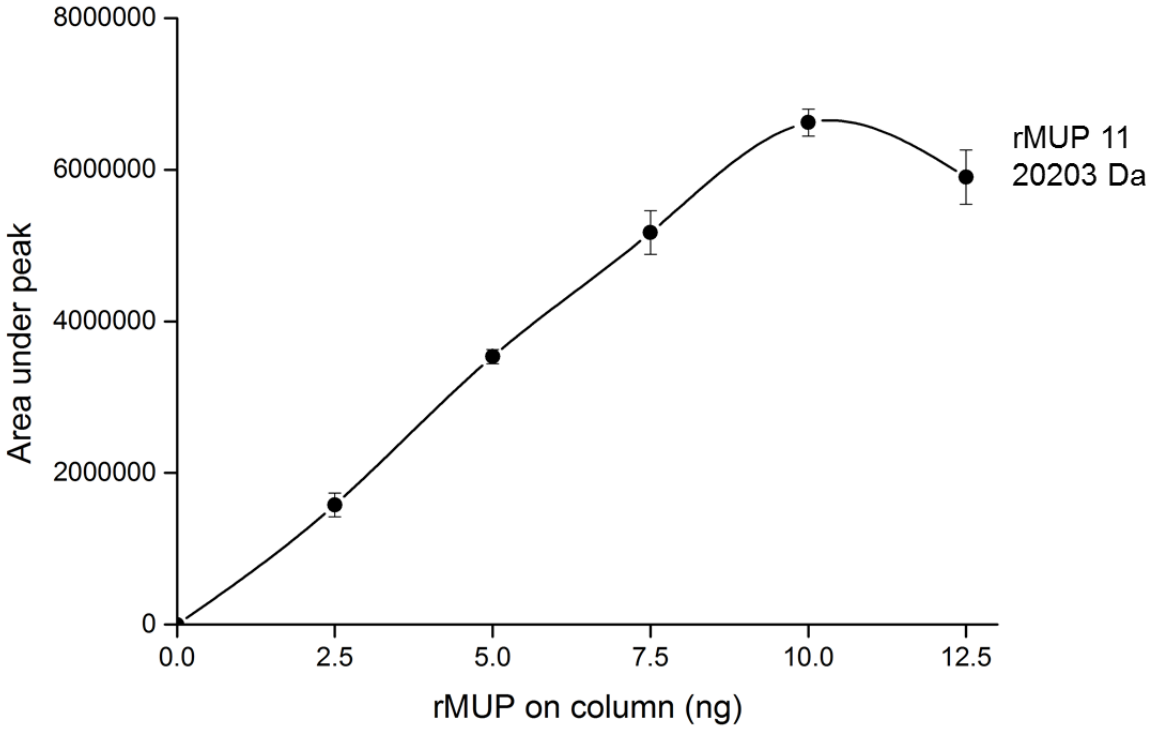
95 percent confidence interval:

 0.2407622 0.2950301

sample estimates:

mean of x mean of y

0.2678961 0.0000000



Supplementary material C: Linearity of ESI-MS intensity with increasing protein load.
rMUP 11 (diluted in 0.1% (v/v) formic acid) was injected onto a C4 desalting trap and the mass of the rMUP was determined by ESI-MS. Data was processed using maximum entropy software MAX ENT1 (MassLynx 4.1, Waters). Sample was injected in triplicate. The area under peak values for rMUP 11 were assessed as total rMUP load on column increased. Error bars represent SD (n = 3).